DOTTORATO DI RICERCA IN BIOLOGIA

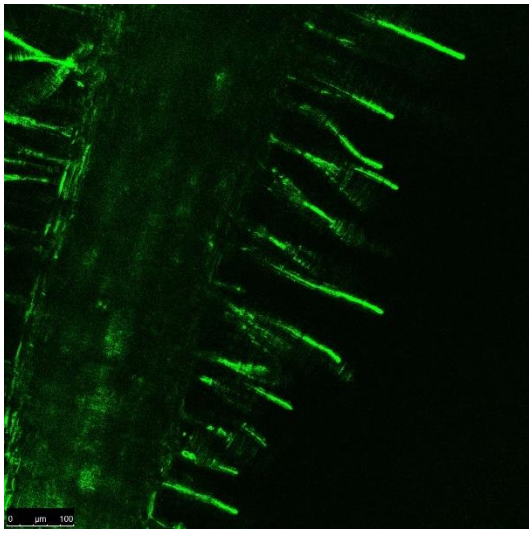
Curriculum Genetica e Microbiologia

(XXIX CICLO)

Plant colonization: exploring the genetic basis of rhizobial symbiotic performance and bacterial invasion

Tesi di

Alice Checcucci



Coordinatore **Prof. Alberto Ugolini**

Tutor **Prof. Alessio Mengoni**

2016

UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA IN BIOLOGIA

Curriculum Genetica e Microbiologia

(XXIX CICLO)

Plant colonization: exploring the genetic basis of rhizobial symbiotic performance and bacterial invasion

Settore Scientifico Disciplinare BIO/18

Dottoranda

Dott.ssa Alice Checcucci

Tutor

Prof. Alessio Mengoni

Coordinatore

Prof. Alberto Ugolini

Anni

2013/2016

Contents

Abstract

Plant growth can be influenced by a large variety of biotic and abiotic factors. Among these factors, the microrganisms associated with plants play an important role. Plant-associated microrganisms (fungi and bacteria) can be present on the surface of plant tissue (rhizospheric and epiphytic), or colonize the internal tissues of the plants (endophytes). Rhizobia represent a paradimantic group of microrganisms which have pecullariar plant-association. They are soil bacteria, which colonize also the rhizosphere of plants, and on some of them (legumes) can enter the root or stem tissues establishing a symbiotic interaction. Under this interaction, rhizobia positively influence the host's growth thanks to the nitrogen fixation process, by which the bacterium is able to fix nitrogen from the atmosphere, making it bioavailable for the plant. The symbiosis is a highly specific process and it is strongly regulated by genetic and chemical factors, which allow the formation of new organs, the root (or stem) nodules. A well-established model used for the nitrogen-fixing symbiosis is that involving the rhizobium *Sinorhizobium meliloti* and the host legume *Medicago sativa* (alfalfa). *S. meliloti* is a particularly interesting study model for its high genetic and phenotypic variability and for the multipartite nature of its genome.

This thesis is focused on three main aspects which characterize *S. meliloti* biology: the interaction and competition for the symbiosis (sociomicrobiology), the relevance of dispensable genes in such interaction, the functional meaning of the multipartite genome.

From the sociomicrobiological point of view, the study mainly focused on the cooperative interactions established among different

strains, which are able to colonize the same plant and, interestingly, the same root nodule. This capability allowed the evolutionary persistence of strains with highly different symbiotic phenotypes.

Concerning the dispensable genome fractions which may affect strain competition, we focused on *acdS*, an accessory gene present in *S. meliloti* strains and in some other rhizobia. This gene encodes the enzyme 1-aminocyclopropane-1-carboxylate (ACC) deaminase, which is postulated to be involved in the decrease of ethylene production by the host plant and in the improvement of the symbiotic performance. Actually, its role in the symbiosis has not been completely elucidated. In our work, we present a phylogenetic and a functional analysis of *acdS* that allowed to hypothesize a role linked to rhizospheric or endophytic colonization.

The last part of the thesis focuses on the bacterial cell metabolic reconstruction, which allowed to deep of the proposed step-by-step evolutionary scenario of *S. meliloti* multipartite genome. The study assessed the contribution of each replicon (which composes the multipartite genome of *S. meliloti*) to its life cycle, in particular in relation with the main three colonized environments (soil, rhizosphere and the root nodule).

Overall, this thesis aimed to design a social, molecular and metabolic profile of *S. meliloti*, trying to improve its application in agriculture and biotechnology and to understand the ecology and the evolution of plant-microbe symbiotic interaction.

Riassunto

La crescita delle piante può essere influenzata da una grande varietà di fattori biotici e abiotici. Tra questi, i batteri che abitano suolo e rizosfera hanno un ruolo fondamentale, tanto che negli ultimi anni il loro uso come biofertilizzati o agenti patoprotettivi è aumentato enormemente. I batteri capaci di colonizzare i tessuti interni della pianta sono chiamati endofiti; tra questi un particolare gruppo è rappresentato dai rizobi, che sono geneticamente programmati per infettare le radici delle leguminose, stabilendo con esse una interazione di tipo simbiotico. I rizobi influenzano positivamente la crescita della loro pianta ospite grazie al processo di azotofissazione, tramite il quale il batterio simbionte fissa l'azoto atmosferico rendendolo biodisponibile per la pianta. La simbiosi è un processo altamente specifico e fortemente regolato da fattori genetici e chimici da parte di entrambi i partner, e che permette la formazione di nuovi organi radicali, i noduli. Il modello biologico utilizzato per lo studio della simbiosi azotofissatrice è quello che coinvolge *Sinorhizobium meliloti* e la leguminosa *Medicago sativa*. *S. meliloti* è un modello studio particolarmente interessante in quanto caratterizzato da un'alta variabilità genetica, fenotipica e dalla natura multipartita del genoma.

Questa tesi si è quindi soffermata su tre aspetti principali che caratterizzano la biologia di *S. meliloti*: le interazioni e la competizione nella simbiosi (gli aspetti sociomicrobiologici), l'importanza della frazione accessoria del genoma in queste interazioni (aspetto molecolare) e il ruolo funzionale ed evolutivo del genoma multipartito.

Dal punto di vista sociomicrobiologico, lo studio si è focalizzato principalmente sulle interazioni di tipo cooperativo che si possono stabilire tra ceppi diversi che possono colonizzare lo stesso nodulo radicale, e che permettono una persistenza evolutiva di ceppi con fenotipi simbiotici altamente diversi tra loro.

L'aspetto molecolare su cui ci siamo concentrati è invece legato al genoma accessorio che può influenzare la competizione tra ceppi. Ci siamo infatti soffermati sullo studio di *acdS*, un gene accessorio presente sia nei ceppi di *S. meliloti* che in altri rizobi. Codificando per l'enzima 1-aminocyclopropane-1-carboxylate (ACC) deaminasi, studi precedenti hanno mostrato un suo coinvolgimento nella diminuzione della produzione di etilene da parte della pianta ospite e nel miglioramento delle performance simbiotiche. In realtà, il suo ruolo nella simbiosi non è stato totalmente chiarito. Nel nostro lavoro, presentiamo un'analisi filogenetica e funzionale di *acdS* che ci ha permesso di ipotizzarne l'origine polifiletica e un coinvolgimento nella colonizzazione rizosferica.

L'ultima parte della tesi si è invece focalizzata sulla ricostruzione del metabolismo cellulare di *S. meliloti*, che ha permesso di approfondire lo scenario evolutivo attualmente proposto per il suo genoma multipartito. Lo studio ha permesso di valutare il contributo di ciascun replicone (che compone il genoma multipartito di *S. meliloti*) al suo ciclo vitale, in particolare, in relazione alle tre nicchie ecologiche che può occupare (suolo, rizosfera e nodulo simbiotico).

Nel complesso, questa tesi si propone di delineare un profilo delle caratteristiche di interazione sociale, dei meccanismi molecolari e metabolici del batterio simbionte *S. meliloti*, mirando a migliorare

la sua applicabilità in campo agricolo e biotecnologico e proponendosi di approfondire il suo ruolo ecologico e l'evoluzione delle interazioni simbiotiche pianta- batterio.

List of Tables

Chapter 1 Tables

Appendix Tables

Section 6.2 Tables

List of Figures

Chapter 1 Figures

Appendix Figures

6.2 Section Figures

Chapter 1- Introduction

1.1 A panoply of plant-bacteria interactions

1.1.1 PGPB: functioning and applications

A large variety of factors can influence plant growth. These can be abiotic or biotic, as the environmental conditions, which include temperature and water, the chemical and physical characteristics of the soil (including the contamination by heavy metal and toxic compounds), the bioavailability of nutrients and the pathogens which can negatively infect the plant (Singh et al., 2015). All these factors can strongly influence the agricultural productivity worldwide. Furthermore, the need for an improvement of the agricultural techniques to cope these stresses, is flanked by the necessity of a more sustainable and environmentally friendly practices (Glick, 2012).

In this purpose, in the last years, the use of microbial organisms instead of chemical pesticides and biofertilizers is getting preferred. As it is known, the rhizosphere is the richer environment in terms of bacteria with beneficial influence on the plant growth (referred as PGPB- Plant-Growth Promoting Bacteria) (Bashan and Holguin, 1998). In literature, their capability to protect the plants from abiotic stresses and pathogens and to stimulate their growth is highly documented (Reed and Glick, 2013). Indeed, they can reduce the susceptibility to diseases caused by pathogens by competition or by the production of particular compounds for the suppression of pathogenic infections (antagonism). In addition, they can increase

the bioavailability of nutritional sources (with nitrogen fixation, phosphate solubilization and iron sequestration), produce various compounds as phytohormones for the root development (Gamalero and Glick, 2015) and contribute to the modulation of the plant defense mechanism referred as induce system resistance (ISR) (van Loon et al., 1998).

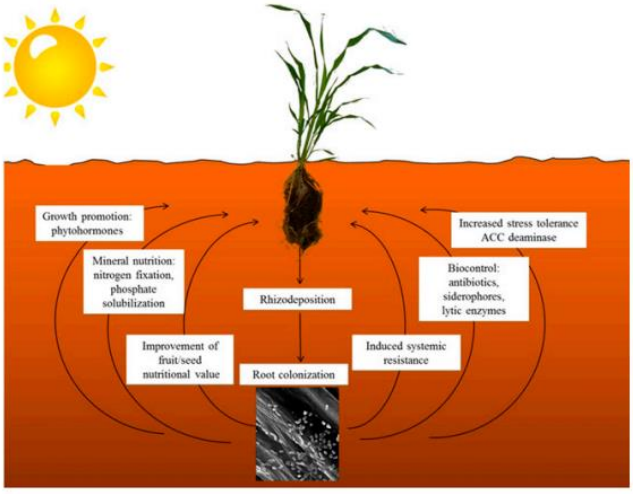


Figure 1.1 Schematic overview of the main mechanisms used by PGPB. From (Gamalero and Glick, 2015).

1.1.2 The endophytic colonization

Thanks to the nutrient-rich exudates released by the plant roots, a large variety of bacteria can establish in the rhizosphere and, some of these, can gradually penetrate the root tissue and choose their lifestyle inside the host (Gamalero and Glick, 2015). The endophytes (both fungi and bacteria) are usually defined as "microorganisms that colonize healthy plant tissue without causing

obvious symptoms or produce obvious injuries to the host" (Bacon and Hinton, 2006;Reinhold-Hurek and Hurek, 2011). The endophytic association and way of colonization is not totally understood, but it is known that their infection is not organ-specific, because a large variety of endophyte were isolates from different plant tissues (Rosenblueth and Martinez-Romero, 2006).

Before the colonization of the plant internal tissues, the endophytes colonize the rhizosphere or the phyllosphere (Sturz and Nowak, 2000). They can enter inside plant through different mechanisms, as tissues wound, via roots hair or with the direct production of degradative enzymes (Reinhold-Hurek and Hurek, 2011). Once penetrated inside the host, the endophyte may localize itself at the point of entry or may diffuse among the plant through vascular system or intercellular spaces.

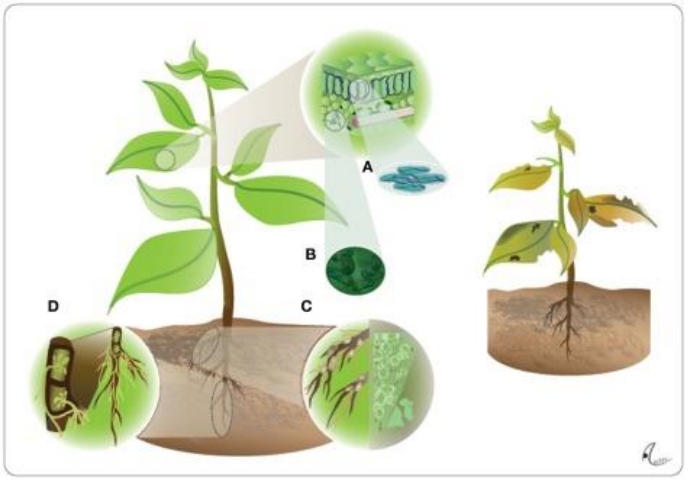


Figure 1.2 Different endophytes can colonize all the plant organs. Endophytes including fungal leaf endophytes (A), bacterial leaf endophytes (B), Rhizobia (C), and arbuscular mycorrhizal fungi (D), can infect all organs of a plant. From (Partida-Martinez and Heil, 2011).

It is important to know that the true endophytic lifestyle can be only confirmed by the presence of the microorganism inside the plant for long time after the colonization, and the endophytic state must be maintained for generation among different plant organs. This could mean a very strong and deep association between plant and its associated bacteria during evolution.

As previously reported, when their presence influence positively the plant growth and health, the bacteria can be classified as PGPB. Bacterial endophyte can serve their host efficiently under a larger range of environmental conditions. It is clearly demonstrated that in many cases they have a protective function for the plant against foreign phytopathogens, nematodes or insects infections (Alvin et al., 2014) through active hydrolytic enzymes or active molecules (Berg, 2005;Berg and Hallmann, 2006). In this purpose, in the last decades, the interest on the potential use of these compounds in the agrochemical, biotechnological and pharmaceutical fields increased a lot (Christina et al., 2013).

Rhizobia are soil bacteria which can lives in a large variety of habitat. Besides the inhabiting of the soil as free living organisms, they can represent a particular group of endophytic bacteria genetically programmed to infect and colonize plant roots establishing a mutualistic symbiosis with specific leguminous plants. Their capability to improve the plant growth is primarily based on the nitrogen fixation process (more detail in section 1.3.3). This highly specific and highly regulated interaction is able to induce genetic, physiological and morphological modification in the legumes, which leaves to hypostasize a long history of co-evolution.

1.2 Legume-rhizobia symbiosis co-evolution

1.2.1 Evolution of the *Leguminosae* (*Fabaceae*) family and its associated rhizobia

Despite the difficult fossils dateability, recent phylogenetic analyses based on molecular data have allowed to characterize past occurrence and habitat distribution of the first leguminous plants (Lavin M. et al., 2005). The first legumes ancestor seems to originate in the late Cretaceous, approximately around 70-60 million years ago in Africa. However, during the Eocene (mid Tertiary, 34-56 Mya) Africa, Europe and North America were still geographically one continuous land, therefore amore recent and detailed hypothesis places the original distribution of legumes in the region of the north Thetys Sea, a strip of salt water that separated the two enormous ancient lands Gondwana (South America, Africa and Australia) and Laurasia (North America, Asia and Europe) (Sprent, 2007a). Here, the conditions were semiarid, which explain the first succulent legumes biome, with a progressive spread in other biomes when the climate change towards cooler temperature (Schrire B.D. et al., 2005).

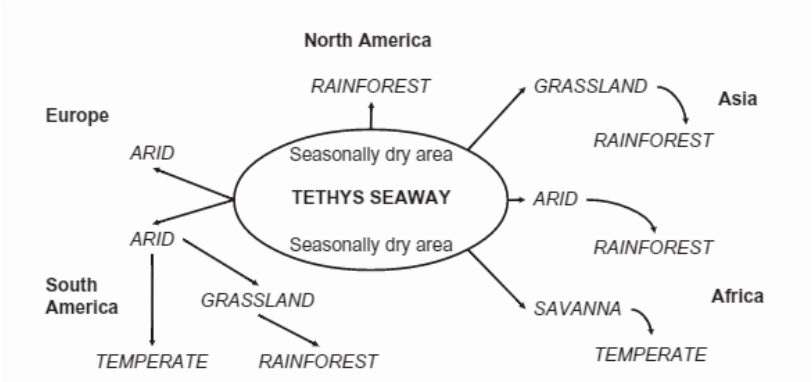


Figure 1.3 Suggested movement of legumes from seasonally dry areas either side of the Tethys Seaway about 60 million years ago to their present locations. After Schrire et al. (2005) from (Sprent and Gehlot, 2010).

The divergence between the three subfamilies (*Caesalpinioideae*, *Mimosoideae* and *Papilionoideae*) of the *Fabaceae* seems to have occurred about 50 Mya, after the spreading and worldwide diffusion of the *Fabaceae*.

The *Fabaceae* family is the third largest family of angiosperm. Presently, the members of *Fabaceae* are distributed in different ecological habitats, with different temperature and humidity condition, and can be annual or perennial herbs and shrubby tree. For many years, a progressive and orderly diversification among the three subfamilies was hypothesized, especially because nodulation and bacterial infection processes are not totally shared among them. However, recent molecular and phylogenetical studies (Lavin M. et al., 2005) award the nodulating ability and the occurrence of plant-bacterial interactions to members of all legumes subfamilies. Actually, the symbiotic interaction with rhizobia with the formation of root or stem nodules (nodulation) is rare in *Caesalpinioideae*, more common in *Mimosoideae* and very common in

Papilionoideae. This pattern of distribution of the nodulation seems to support the temporal order in which these subfamilies evolved (Allen and Allen, 1981).

However, as previously mentioned, the first legumes fossils are not easily identifiable, mostly those with nodule infections. Thus, it is not totally clarified how long ago the first legumes started to associate with rhizobia.

Considering the complete absence of rhizobial fossils evidence (where rhizobia stand for "bacteria with the ability of nodulation and nitrogen fixation"), it is difficult to date their speciation, especially in relation with the host plant. Some hypotheses (Sprent, 1994) described the possibility of two separate nodulation events occurred in the late Cretaceous, one that involved the ancestor of *Rhizobium*, the other one that involved the slow growing *Bradyrhizobium*. Furthermore, it is possible that the rhizobia-legumes symbiosis is originated from more general pathogenic interactions. More recently, a phylogenetic analysis based on a set of conserved genes in different genera of rhizobia allowed evidence that the appearance of rhizobia anticipated the arise of their host leguminous plants (Turner and Young, 2000). Consequently, rhizobia, originated before of their host plants, likely as free-living soil or rhizospheric bacteria.

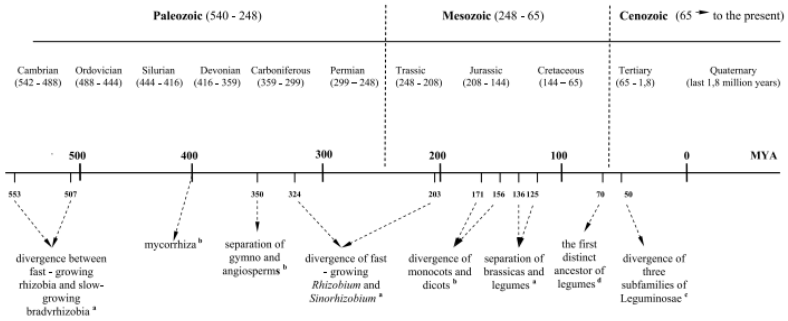


Figure 1.4 Estimated divergence times in rhizobium-legume symbiosis. From (Dresler-Nurmi et al., 2007).

1.2.2 Nodulation: evolution and revolution

In addition to wonder when and where legumes evolved, it is interesting to ask why rhizobial-plant association evolved in some of them. It is important to take in mind that the evolution of this association is not completely independent. As rhizobia, fungi can establish symbiotic interaction with the roots of leguminous (and not leguminous) plants. In the mycorrhizal association, the fungus colonizes the host plant's roots, both intracellularly, as in arbuscular mycorrhizas (AMs), and extracellularly, as in ectomycorrhizas (ECMs). As well as rhizobia, they represent an important component of soil and of its chemical balance, enhancing nutrient uptake and parasitic defense. The AMs and the ECMs in legumes evolved long before them (Sprent and James, 2007), assuming that, potentially, all legumes can produce mycorrhizas following the fungus symbiotic association. This is an ancient symbiosis verifiable in a large variety of plant species, which has been detected in fossils of early land plants more than 400 million years ago (Stracke et al., 2002). Several studies showed the evidence of a common

evolutionary development of the molecular players of rhizobial and fungal association with their respective legume hosts, sharing some common key genes. In particular, the two symbiotic associations overlap genetic programs deeply related in terms of common the signaling pathways. In both the host model species used for the associations study (*Medicago truncatula* and *Lotus japonicas*), the common symbiotic pathways comprise a conserved set of genes encoding a membrane receptor-like, a nuclear cation- channel and several interacting proteins important for fungal and bacterial recognition (Stracke et al., 2002;Streng et al., 2011). Thus, in legumes the common signaling pathway is activated by distinct Lys-M type receptors kinase highly specific for fungal or bacterial signal molecules (respectively, Myc factor and Nod factor).

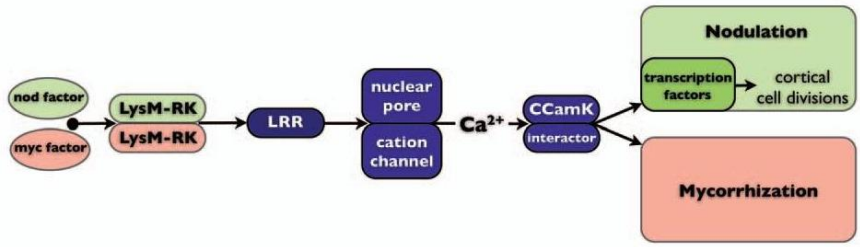


Figure 1.5 Schematic representation of the genetically dissected symbiosis signaling pathway in legumes from (Streng et al., 2011).

Legumes can be very versatile in their symbiosis, consequently, multiple concomitant symbiotic association coud have taken place. Indeed, nodulation has a significant requirement for phosphorus, so it can be inclined to establish P-acquiring symbioses as AMs and ECM near nodules (Sprent and James, 2007). In contrast to the

mycorrhizal association, the nitrogen fixing symbiosis evolved more recently, and the host and partner choice is more phylogenetically restricted. Several studies confirmed that the following few million years after the first legumes appearance (55 Mya), were characterized by a great climate change. Indeed, there was a strong increase of temperature, which favored the carbon mobilization from sea-sediment and its cycle through terrestrial ecosystem. It is known that the nitrogen fixation process uses a large amount of the total carbon produced by the plant. Then, it was hypothesized that the high level of carbon dioxide in the atmosphere, coupled to a decrease of bioavailable nitrogen, favored the evolution of symbiotic nodulation (Bowen et al., 2004;Sprent, 2007b;Sprent and James, 2007).

The genomic rearrangement and the horizontal gene transfers (HGT) among different species are the most important factors in bacterial evolution (Gogarten et al., 2002;Koonin, 2009). Molecular analyses upon the rhizobial symbiotic accessory genes allowed to establish a quite ancient origin of them (Moulin et al., 2001). Possibly, they evolved to their actual form thanks to divergence and transfer events among different genera of α- and β- *Proteobacteria* (both the orders comprise rhizobia) and strains from different geographical areas (Provorov, 1998). Under these hypotheses, it was proposed that the common ancestor of rhizobia was a free-living organism with the additional capability to fix nitrogen, which was preserved during evolution only by nodulating strains (Dresler-Nurmi et al., 2007).

The evidence of bacterial genetic plasticity may then suggest a strong ability of rhizobia to adapt to different host legumes, more than a direct effect on plant evolution. Furthermore, the correspondence between the phylogenies of some rhizobial genes

(also related to the nodulation) and some genomic elements of the host plants, could suggest a co-evolution model of the symbiotic association (Dobert, 1994).

The coordination that can allow the establishment of a symbiotic interaction, especially in terms of morphological modification and molecular mechanism of regulation, could be explained by a recent model proposed by Werner et al. (Werner et al., 2014a). The study suggests a "predisposition hypothesis", which demonstrate the evidence of a single and necessary innovation event as basis of nitrogen fixation evolution. Indeed, it seems that a single shared symbiotic trait among leguminous plants, defined as "precursor state", could have represented the initial stage of nodulation process, which in some cases evolved as "stable fixer state" following additional mutations but in other, it could also disappear. The proposed single nitrogen fixation origin could represent the explanation for the homology existing between all the symbiotic association established by different rhizobial microorganism and their specific legume host.

1.2.3 Nodules and nodulation types

Nodulation is a complex mechanism evolved by rhizobia which allows the interaction with specific host plants, establishing a continuous and cooperative exchange of benefits between partners. Firstly, rhizobia penetrate into the legume roots. The infection can occur directly through the epidermis or through breaks (referred to cracking infection).

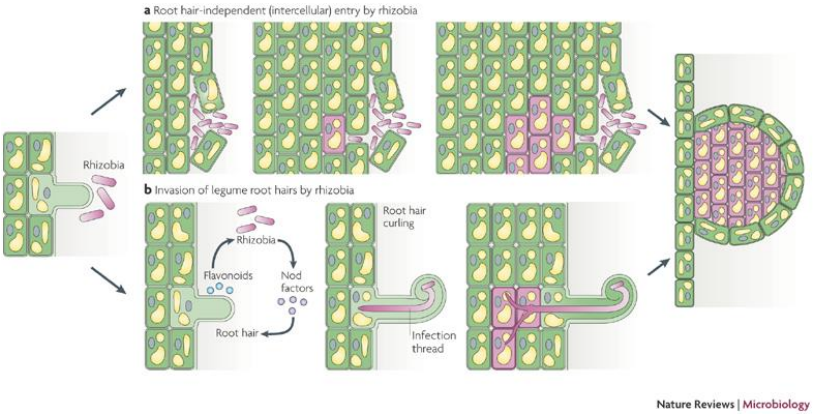


Figure 1.6 Bacterial penetration of the root tissue can occur by two different strategies. In root hair-independent (intercellular) entry (a), disruption of the epidermal cell layer allows rhizobia to enter the root. This contrasts with the classic model of root hair invasion (b), where flavonoids released by legume roots trigger the synthesis of rhizobial Nod factors, which induce root hair curling, bacterial penetration at the center of the infection pocket and division of cortical cells. From (Deakin and Broughton, 2009).

The most studied and diffused root hairs entry is the highly regulated epidermal infection, and involves transcellular colonization. The infection occurs in some but not in all the cells of the root invaded region. These cells divide repeatedly and enlarge their size giving a well know swollen structure to the nodule, usually retaining the meristematic activity (Sprent and James, 2007;Sprent, 2008). However, the plant species that usually are characterized by the epidermal way of infection, could sometimes become infected by the formation of wounds, especially under stress condition. Actually, the breaks can be a normal way of infection which can include the formation of occasional lateral roots. This event can characterize particular plants as dalbergioid and genistoid clades of legumes, which include grain as *Lupinus* and *Arachis*, and forage as

Stylosanthes.(Sprent and James, 2007). It was hypothesizing (Sprent and James, 2007) that both the epidermal and the breaks infections may include the possibility of the formation of two different pattern of nodule development: one that includes the transcellular infection and one not. Then, two type of nodules could evolve from each of these: the determinate one and the indeterminate one.

The mechanism of indeterminate nodules formation is the most common in legumes. This mechanism was favored during evolution probably for the attitude of this nodules to survive after years (indeed, it is very common in perennial legumes) and under stress condition. The indeterminate nodule development includes the formation of the root hair curling, the fast division of root cortical cells and the differentiation of the nodule primordium (see figure 1.6). Inside the new formed structure, the bacteria are surrounded by a membrane and some of them differentiate in the symbiotic form able to fix nitrogen, the bacteroid (for details, see section 1.3). The maintenance of the active apical meristem (which confer to the nodule the typical cylindrical shape) allow to the nodule cells a continuous replication and regeneration, creating a gradient of infected and uninfected cells among the structure. It is possible to detect four different regions, commonly called zones, which demark different stages of the symbiotic process and of the bacteroids differentiation. Indeed, not all bacteria inside the nodule are differentiated into nitrogen fixing bacteroids. Bacteroids are mainly located in the III and IV zones (respectively nitrogen fixation and senescence zone, closer to the root) and a fraction of undifferentiated bacterial cells is maintained in correspondence of the apical tissue.

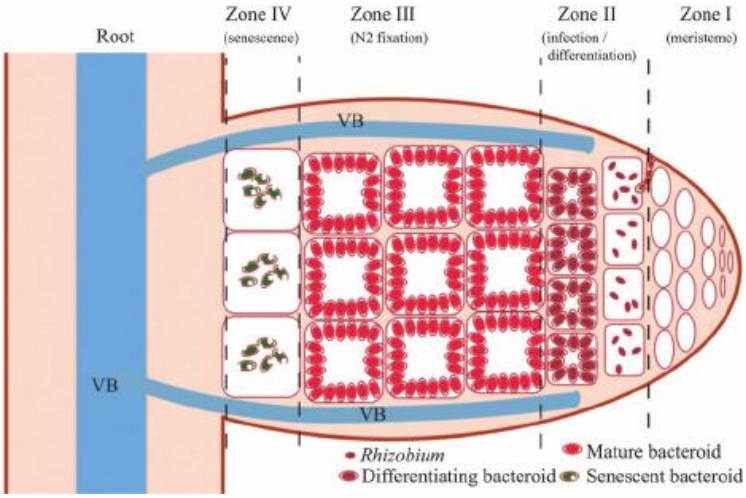


Figure 1.7 Diagram of indeterminate nodule development. The nodule grows by meristematic cell divisions in zone I. In zone II, rhizobia are released by the infection thread into cortical cells by an endocytic process. There, rhizobia differentiate into bacteroids that in zone III synthesize nitrogenase and other enzymes involved in nitrogen fixation. After some weeks, the nodule starts senescing and bacteroids are degraded (zone IV). VB indicates Vascular Bundle. From (Rodríguez-Haas et al., 2013).

Contrarily, in the determinate nodules, the meristematic activity disappears shortly after nodule formation, and the fast growth of cells after the infection resulting in a spherical shape of the nodule. Here, all the bacteria assume the form of bacteroid, able to fix nitrogen, but not terminally differentiate.

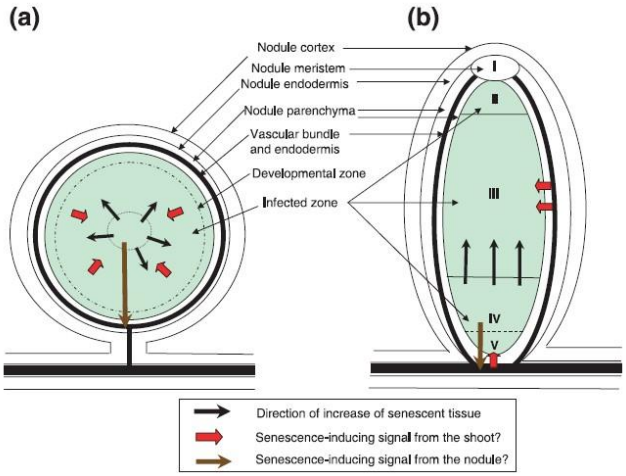


Figure 1.8 Comparison between the structure of determinate (a) and indeterminate nodules (b). Diagrammatic representation of the components. From (Puppo et al., 2005).

If a way of infection directly corresponds to a particular nodule type, is still not totally clarified (Sprent, 2007a). However, it is recurrent that the determinate nodules preferentially originate following a crack infection. On the contrary, the indeterminate nodule formation seems to be more common for the epidermal infections.

In the last decades, evolutionary studies (Sprent, 1980) have hypothesized the origin of determinate nodule from the indeterminate nodule, following precise metabolic requests. The rapid development and the fast cell division which characterize the determinates, can be evolved following an environment adaptation, possibly to short life plants as the annual crop and the shrubby species (Sprent, 2001). On the contrary, the indeterminate nodules are able to endure long period of stress. This could be related also to

host plants which colonizes more temperate and seasonal areas (Sprent, 2007a).

1.2.4 What remains of nodules?

Of course, the lifetime of the nodule is extremely variable, and it depends on the legume species and on the environmental and stress conditions to which the host plant is exposed. Then, senescence is a natural stage of the symbiotic nodule life, and it allows the nutrients recirculation among the organisms of the same habitat (Serova and Tsyganov, 2014).

For their typical round-shaped form, determinate nodules are characterized by a radial senescence of the tissue (see figure 1.8). Contrarily, accordingly to the differentiation-zones structure, in the indeterminate nodules the senescence area (IV zone) gradually moves to the apical part initiating the degeneration of the nodule (Van de Velde et al., 2006). Phenotypically, the nodule senescence is characterized by the evident color shift of the tissues from the typical pink (indicating the presence of leghemoglobin in plat tissues, then an active nitrogenase) to green, which is determined by the degradation of the leghemoglobin heme group (for details see section 1.3.3) (Roponen, 1970). Biochemically, senescent nodules are characterized by an intense proteolytic activity that causes a progressive protein degradation (Pladys and Vance, 1993).

If the studies performed on the nodule senescence have been well deepen from the plant physiology and biochemistry, the release of the symbiotic bacteria in the soil is still not totally explored. As previously mentioned, the bacteroids in determinate nodules are

densely distributed and remain quite similar to the free-living rhizobia. Indeed, they are not terminally differentiated, but assume a reversible form. On the contrary, bacteroids in indeterminate nodules are functionally and morphologically distinct from free-living rhizobia. Then, when determinate nodules die and release the bacteroids in the soil, they can easily de-differentiate to the free-living form, enriching the soil populations of substances stored during their symbiotic-life. In particular, the carbon storage polymer poly-β-hydroxybutyrate (PHB) is one of the most important metabolism product of the bacterial cells, as in most of the rhizobial species. The storage of PHB for rhizobia can mostly occur during free-living lifestyle, but it can occur also during the symbiotic process (Trainer and Charles, 2006). It was largely demonstrated that the bacteroids able to accumulate PHB in determinate nodules and in the undifferentiated cells of indeterminates, may have a competitive advantage when released into the soil after nodule senescence (Denison, 2000).

On the contrary, in indeterminate nodules, the terminally differentiated bacteroids die with the nodule, but the still maintained undifferentiated rhizobia are released from the nodule tissue giving rise to a new rhizobial soil population (Denison, 2000).

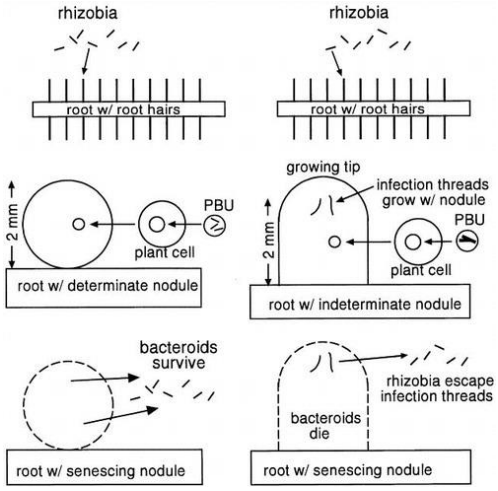


Figure 1.9 Life cycle of rhizobia from soil to soil via determinate (left panel**) or indeterminate** (right panel) **legume root nodules**. PBU (peribacteroid unit). From (Denison, 2000).

1.3 Molecular details of the legume-rhizobia symbiosis

1.3.1 The importance of nitrogen

The nitrogen is the most abundant gas in Earth's atmosphere (78%), and is an important element for the life. The nitrogen could be present in the ecosystems in a large variety of chemical forms, as ammonium (NH_4^+), nitrite (NO_2^-), nitrate (NO_3^-), nitrous oxide (N_2O), nitric oxide (NO) or inorganic nitrogen gas (N_2) (Postgate, 1982). The nitrogen cycle, in which atmospheric nitrogen (N_2) is converted into different available organic compounds, is one the most crucial natural processes to sustain living organisms, in particular the plants. The conversion can be sustained by biological or physical mechanisms, and is primary for its circulation among the ecosystems. Prokaryotes are the only organisms able to utilize molecular nitrogen, and in particular, the symbiosis between bacteria and leguminous plants alone provides 20% of the global biological nitrogen fixed annually (Dresler-Nurmi et al., 2007). For this reason, the use and the control of the activity of this bacteria and the process of Biological Nitrogen Fixation (BNF) could be considered a fundamental element for a sustainable agriculture (Stewart, 1966).

Prokaryotes able to perform the BNF are called diazotrophs, and their fixing behavior can be very variable. One of the most studied BNF, is exactly that which involves the leguminous plants roots and the rhizobia.

1.3.2 Symbiosis: host specificity, infection and nodule formation

A large variety of bacterial species belonging to α- and β-*proteobacteria* are able to perform the nitrogen fixation on leguminous plants (Chen et al., 2003). Among all the BNF, that which involves *Sinorhizobium meliloti* strains and *Medicago*

sativa is one of the most studied. Indeed, *M. sativa* (commonly named alfalfa) is a perennial leguminous plant cultivated as forage crop in many countries of the world. It is native of temperate climate but it is known to be tolerant to a large variety of temperature. *M. sativa* can be used as forage, for the high protein and vitamin content (important for the ruminants' nutrition), but also as biofuel and for soil quality improvement (Mozaffari et al., 2000;Bouton, 2007;McCaslin and Miller, 2007). Indeed, alfalfa cultivation naturally allows to enrich the soil with nitrogen after the depletion due to the previous crops (e.g. cereals) cultivation.

The mechanism which allows the symbiotic interaction between the nitrogen fixing rhizobia and leguminous plants, is highly regulated. Indeed, both the organisms have evolved a complex and specific signals exchange that let to set up the multi-step mutualist association (Jones et al., 2007). In *Sinorhizobium meliloti* (Capela et al., 2006), the symbiotic process starts with the mutual recognition between the plant and the bacteria present in the rhizosphere. Indeed, in general, the soil area close to the roots is exposed to a large variety of microorganisms, and can represent an important source of nutrients for bacteria due to the presence of root exudates. The host legume is able to secrete specific signals represented by flavonoid compounds (2-phenyl-1,4-benzopyrone derivatives, as for instance

luteolin). They selectively bind a receptor protein in the bacterial cell (NodD), activating the expression of the bacterial *nod* genes, which allows the production and secretion of the Nod factor (Shaw et al., 2006), a chito-lipo-oligosaccharide molecule. The Nod factor, in turn, triggers the expression of plant developmental cascade for nodule formation (starting with root hair curling). The exact chemical structure of the Nod factor and the flavonoid secreted by plant, are the basis for host–symbiont specificity (Kondorosi et al., 1984).

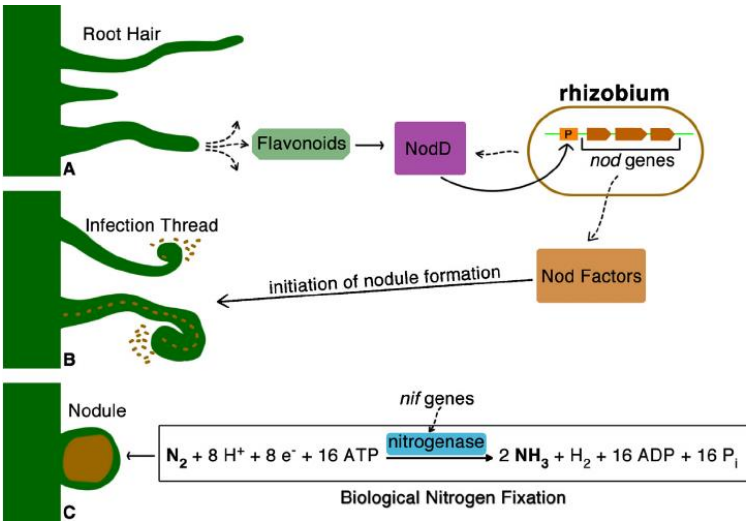


Figure 1.10 Schematic overview of the nodulation process and biological nitrogen fixation From (Laranjo et al., 2014).

The root hair curling and the cortical cell divisions are the following step of the symbiotic process. This step favours the invasion of the plant root cells and the formation of the infection thread by bacteria (Shaw et al., 2006). When the rhizobia are completely internalized

by root cells, they are included by endocytosis in vesicles named symbiosomes, where they differentiate in specialized forms, the bacteroids (Brewin, 1991;Brewin, 2004). In the same time, the infected root cells undergo an enlargement and produce a swell on the root surface. This process led ultimately to the formation of a new organ, the root nodule. Inside the microaerobic environment of the nodules, the bacteroids expresses the oxygen-sensitive enzyme nitrogenase that catalyzes the conversion of atmospheric nitrogen to ammonia (Peters and Szilagyi, 2006).

1.3.3 Nitrogen fixation in legume-rhizobia symbiosis

The nitrogen fixing enzyme complex, named nitrogenase, consists of two protein components, called the Fe-protein and the MoFe-protein. Both are required for the nitrogen fixation process, that consist in a step by step protein association and dissociation and the hydrolysis of ATP molecules for the delivery of the electrons (Peters et al., 1995). The nitrogenase complex is produced by certain prokaryotes, including species of *Rhizobium* genus, the free-living *Azotobacter*, Actinobacteria belonging to *Frankia* genus, Cyanobacteria (blu-green algae), and in some Archaea (Raymond et al., 2004). In the legumes-rhizobia symbiosis, the nitrogen fixation takes place inside the root nodules. Indeed, it is here that the bacterial cells morphologically differentiate in bacteroids. As previously mentioned, they are characterized by a swollen and irregular form and exhibit a strong downregulation of many classical metabolic function (including the DNA replication) in favor of the activation of the genes related to the nitrogen fixation process (Barnett and Fisher, 2006). Indeed, bacteroids use the enzyme

nitrogenase to catalyze the reduction of dinitrogen with this chemical reaction:

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

The BNF requires a large amount of energy (for every nitrogen conversion, 16 ATP molecules) which is provided by the host plant in the form of photosynthates, basically composed by carbohydrates. These substances are transported to the nodule in the form of sucrose via phloem. Essentially, in the nodule, the sucrose is cleaved by sucrose synthase to UDP-glucose and fructose. Then, the hydrolyzed products are metabolized by glycolytic enzymes to produce phosphoenolpyruvate (PEP), which is finally converted in malate. Here, the dicarboxylates are transported across the symbiotic membrane inside the bacteroid, thanks to the dicarboxylic acid transport (Dct) system, and metabolized via the tricarboxylic acid (TCA) cycle required to fuel nitrogen fixation. Ammonium is the primary stable product of the reaction, it is crucial for the N assimilation for the plant, and it is transferred by diffusion through bacterial cytoplasm to the plant (Lodwig and Poole, 2003).

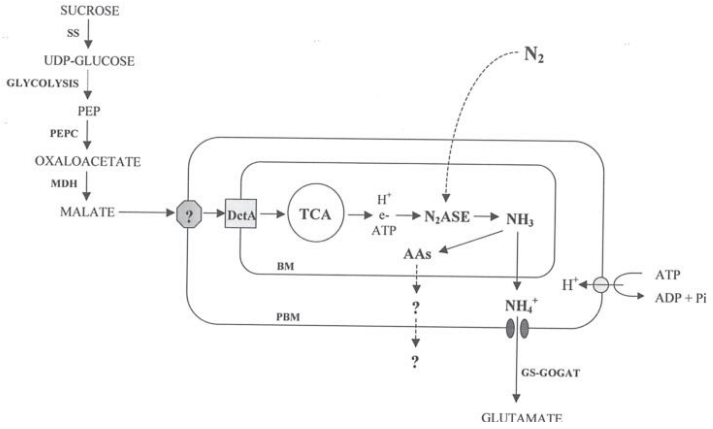


Figure 1.11 Schematic representation of nutrient exchange between a nitrogen fixing bacteroid and the plant cell; AAs, amino acids; BM, bacteroid membrane; GSGOGAT, Glutamine synthetase-Glutamate synthase; MDH, malate dehydrogenase; PBM, peribacteroid membrane; PEP, phosphoenolpyruvate; PEPC phosphoenolpyruvate carboxylase; SS, sucrose synthase; TCA, tricarboxylic acid cycle from (Lodwig and Poole, 2003).

Besides the inclusion in aminoacids as alanine and aspartate, the nitrogen is mainly assimilated into glutamine and glutamate through the action of glutamine synthetase (GS) and glutamate synthase (GOGAT). Fixed nitrogen is then 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. Tropical legumes, such as soybean and *Phaseolus* bean, which form determinate nodules, have a complex biosynthetic pathway for synthesis and export of ureides such as allantoin and allantoic acid (Atkins, 2000).

Asides from the high amount of energy required, the BNF is characterized by some limiting factors. One of this is the amount of oxygen concentration. O_2 can indeed inactivate the nitrogenase

activity. Otherwise oxygen is required by aerobic nitrogen-fixers to provide the large amount of ATP required. Therefore, the bacteroids are able to regulate the amount of oxygen present in the nodules thanks to the leghemoglobin (legume-hemoglobin).

Figure 1.12 Nodules with active BNF. Indeterminate nodules characterized by the typical pink color of active nitrogen fixation from a *S. meliloti- M. sativa* experiment.

This oxygen-binding protein is found in very high concentration inside root nodules. It is composed by an heme group produced by rhizobial cells, and a globinic part provided by the plant. Thanks to its chemical structure, leghemoglobin is able to bind free-oxygen in the nodule facilitating the respiration of the bacteroids and, at the same time, preventing the inactivation of nitrogenase (O'Brian et al., 1987).

1.4 Sociomicrobiology of the legumes-rhizobia symbiosis

1.4.1 Interaction strategies

Since Darwin's studies, the importance of biotic interactions in organisms' evolution, has stirred the attention of a large plethora of biologists. In particular, evolutionary biologists started to develop theories on the evolution of cooperation and its effect in terms of fitness benefits (Wilson, 2000). The interaction mechanisms, according to which individuals in a particular environment can interact with multiple partners, can determine variable costs and benefits, then allowing to evolve new organization types (colonies, multicellularity, etc.), which derive from the interaction among single cells or organisms.

Classically, the interaction strategies can provide benefits for one of the organism involved in the interaction, or can provide a common advantage which allows an increase of fitness for both (or for all, if more than two) the individuals (Werner et al., 2014b). Commonly, the close relationship establishing between two organism of different species is named symbiosis. Regarding the benefit that such organism may confer, three type of association are defined: the mutualism, the parasitism and the commensalism, and they basically differ from each other from the overall shared benefits.

In the mutualism both the organisms involved in the interaction receives benefit from the other, in commensalism one species benefits while the other is unaffected, and in parasitism only one species benefits from the interaction, the other is affected negatively.

Table 1. Schematic summary of the interaction strategies.

Interaction	Species A	Species B
Mutualism	receives benefit	receives benfit
Commensalism	receives benefit	not affected
Parasitism	receives benefit	harmed

The symbiotic interactions can be classified according to the location of the symbiont with respect to the host cells: ectosymbiosis or endosymbiosis. Despite the evolutionary preservation of all three relationship strategies, detailed researches revealed that much of the achieved global diversity is sustained by mutualist-cooperative interactions among different species (Bascompte and Jordano, 2007). Actually, it is known that every species is involved directly or indirectly in mutualistic interactions with one or more partners, and most of these are important for the reproduction and the survival, as the pollination (Potts et al., 2010), the seeds dispersal (Potts et al., 2010) and critically participating to nutrient cycling, as for plant associated with fungal mutualists (Wilson et al., 2009;Kiers et al., 2010). The cooperative interactions between individual of different species and the shared increase of the fitness can be successfully explained as a biological market (Werner et al., 2014b). Indeed, the theory recently proposed, allows to explain the maintenance and the evolution of the mutualism interdependence, considering the costs and the risks for the partners, the active choose of cooperator and the competitions between individuals. Recently, microorganisms have become popular models for addressing sociobiological questions related to the evolution of multicellularity and to the exchange of nutrients. Indeed, microbial system are suitable to study the mutualistic interactions for the readiness of manipulation, the simple tracking of the resources and the short

generation times. In particular, keeping on the parallelism between mutualistic interaction and biological markets, the collaborative behavior established by microbe can be compared to a trading partner. Besides the classical exchange of benefits and the reciprocal control of the fitness, the microbial mutualist evolved the possibility of the partner discrimination, basing on its quality and cooperativeness, and to respond to resource variability, influenced by biotic and abiotic conditions.

The bacterial kingdom can be involved in a large range of beneficial or not-beneficial interactions with different organisms, such animals, fungi and plants. Today, many example of microbial symbiosis are known, and the phylogenetic distribution of symbiotic interactions between prokaryotes and the eukaryotic hosts is much wider than previously thought (Moya et al., 2008).

Many unicellular eukaryotes are able to ingest bacteria through phagocytosis. This can be the case of the ameba *Amoeba proteus*, which can be infected by X-bacteria establishing a parasitic interaction. Interestingly, today we have evidences that, in some cases, this interaction can become a mutualistic association, that makes the bacteria a required cell component for the eukaryote. *Dictyostelium discoideum*, commonly nicknamed "social amoeba", can be used as model for the different interactions establishing with various bacterial classes. This amoeba is indeed able to eat bacteria and human pathogens, or being destroyed by them. (Cosson and Soldati, 2008). The model system used to describe a predator and collaborative social behavior (multicellular formation necessity) is the bacterium *Mixococcus xantus*, which is able to predate other microbes in normal environmental condition, or cooperate forming multicellular structure if in nutrients lack (Berleman et al., 2008).

Advances in genomics have had a large impact on the research into microbial symbioses, especially those with eukaryotes, which has implications for biotechnology (for example, sponge-associated prokaryotes (Taylor et al., 2007), biomedicine (for example, the human intestine microbiome (Stingl et al., 2005) and agriculture (nitrogen fixation in plant-associated bacteria (Kneip et al., 2007). In this case, that which involves rhizobia and leguminous plants is one of the most studied.

This is a cooperative interaction known as beneficial mutualism. It is interesting to know that different rhizobia strains belonging to a single species can behave differently towards the host. Indeed, possibly as a result of multiple mutation (in the bacterial genes correlated with the nitrogen fixation), some rhizobia are able to infect and nodulate the plant, but fixing little or no N_2. These rhizobia are termed "ineffective" strains and establish with the plant a parasitic interaction.

1.4.2 Evolutionary balance is maintained by host sanctions

It is plausible that mutualistic rhizobia easily persisted during the evolution displacing the non-fixing strains, because of the reciprocal beneficial relationships established with the plant. However, nitrogen fixation is an energetically expensive process, which uses resources that rhizobia may utilize for their own growth. Consequently, in theory we can hypothesize that the parasites (not nitrogen fixers) could gain fitness over the mutualists (nitrogen-fixers). A possible explanation for the persistence of the fixers could be the sanctions inflicted by the host plant on the parasites, thus

indirectly favoring the beneficial infection of the mutualists (Akçay and Simms, 2011). According to the sanctions hypothesis (Kiers et al., 2006), the host plant is able to discriminate among nodules according to their capability in the nitrogen fixation. Indeed, the presence of sanctions could support the infections by mutualist strains and inhibit that by the non-mutualist. Therefore, under this hypothesis the host would perform a "partner choice". Different type of symbiotic interactions may be also present in the same host plant, which then, is able to maintain a balance between multiple symbiont genotypes (Kiers et al., 2003). In this case, the sanctions theory seems to be particularly important to describe the stabilization of cooperation

between different strains, which can colonize the same host plant, and, apparently, the same symbiotic nodule. According to the a recent view of rhizobial symbiotic interaction, which uses the theoretical framework of the market economy to explain the benefits of the symbiosis (Werner et al., 2014b), the legume sanctions (which punish rhizobia that fail to fix nitrogen) can warrant the "fair trade" and can ensure the evolutionary persistence of different symbiotic performances.

1.5 *Sinorhizobium meliloti*: a symbiosis study model

1.5.1 Genes for symbiosis

Basically, every bacterial species is characterized by a high variability among the different strains. The concept of pan-genome has been developed to consider these differences (Medini et al., 2005;Tettelin et al., 2008). Considering that new genes are always discovered when a new genome is sequenced, the pan-genome is composed by a core and a dispensable fraction. The core contains genes shared among closely related species (Young et al., 2006), and the dispensable comprises more specialized unique and accessory genes. Therefore, the house-keeping genes are included in the core genome of a species, and are mainly collocated on the principal chromosome. On the contrary, the dispensable genome contains genes more related with the environmental adaptation and survival in particular ecological niches, where the unique genes harbour more specific functions for a single strains (Galardini et al., 2013b). These fraction is usually collocated on the secondary plasmids. Indeed, in the most of rhizobial species, the majority of the symbiosis- related genes, which are not essential for the life of the bacterium, map on plasmids or chromosomal islands (Becker, 2007). The case of *S. meliloti* 1021 is not different from this generalization.

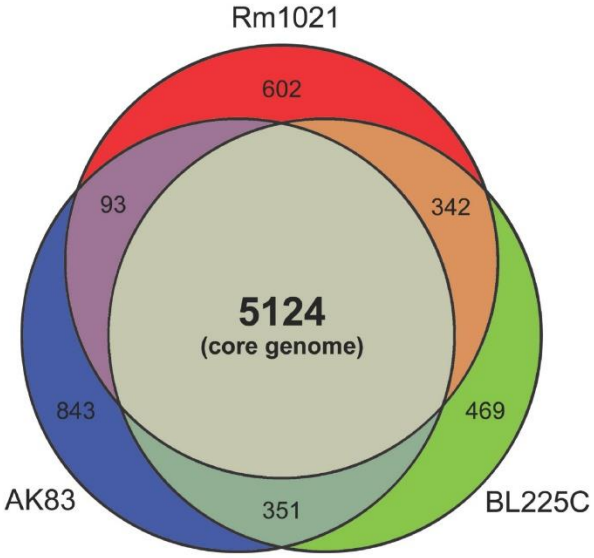


Figure 1.13 Size of core and accessory genome of three *S. meliloti* strains. The number of orthologous groups found in each intersection is reported. Areas are not in scale. From (Galardini et al., 2011).

1.5.2 Comparative genomics and genome evolution of *S. meliloti*

S. meliloti is a gram-negative α-proteobacterium which belongs to the order of Rhizobiales. As other rhizobia, can be found in many soil type as free living organism (Roumiantseva et al., 2002) or in association with specific leguminous plants, on which it induces the formation of nodules. *Medicago sativa* L. (alfalfa) and *Medicago truncatula* are the most studied host species for the *S. meliloti* symbiosis (Gibson et al., 2008). In the recent years, complete genome sequences of a number of rhizobia have been determined, in particular that of the strain *S. meliloti* 1021 (Galibert et al., 2001) has aroused a great interest. The new knowledge based on the

genomic sequences allowed to deeper the molecular details and the major steps of this symbiotic interaction and rhizobial lifestyle.

The genome of several rhizobia is characterized by a multipartite genome structure. In particular, (Galibert et al., 2001) the genome of *S. meliloti* 1021, comprises a primary chromosome (3.65 Mb) and two small units termed megaplasmid pSymA (1.35 Mb) and chromid pSymB (1.68 Mb) (Galibert et al., 2001). This multipartite nature is also present in all the other members of this species sequenced so far (Galardini et al., 2011;Schneiker-Bekel et al., 2011;Galardini et al., 2013a;Galardini et al., 2013b), which may also contain some additional smaller plasmids. The chromosome is the largest replicon, and contains most of the core genes. The pSymB chromid is mainly harbouring genes related to transport and catabolism of carbon sources. The pSymB chromid contains also some essential genes (Harrison et al., 2010), which makes sense of its chromid nature (Harrison et al., 2010)). Finally, the pSymA megaplasmid mostly contains the genes related to the symbiosis interaction and to the nitrogen-fixation process. A large fraction of the genome variability (as genomic rearrangements and presence of the dispensable genome) is located on pSymA (Galardini et al., 2011). pSymA is hypothesized to be the most recent replicon in the *S. meliloti* genome, likely acquired through horizontal gene transfer (HGT) (Wong and Golding, 2003). One of the main aspect of *Sinorhizobium meliloti* pangenome is the high intraspecific variability among the different strains. Only recently, the scientists tried to explore the correlation between this large genetic variation and the symbiotic performances with the host plant (Pini et al., 2011). It was clearly demonstrated that the determinant traits related to the symbiotic capabilities located on the megaplasmid, (as *nod*,

nif and *fix* genes (Batut et al., 1985;Fischer, 1994) are liable to variations. In particular, the genes involved in microaerophilic growth during bacteroid development, those who determines the competitiveness inside nodules, their number for each plant, and those for the nitrogen metabolism can be in some strains completely absent or downregulated (Pini et al., 2011). This can of course determine a variation in the symbiotic phenotype of the bacterium, with an effect on the plant growth, giving a genetic base to the existence of cheater strains.

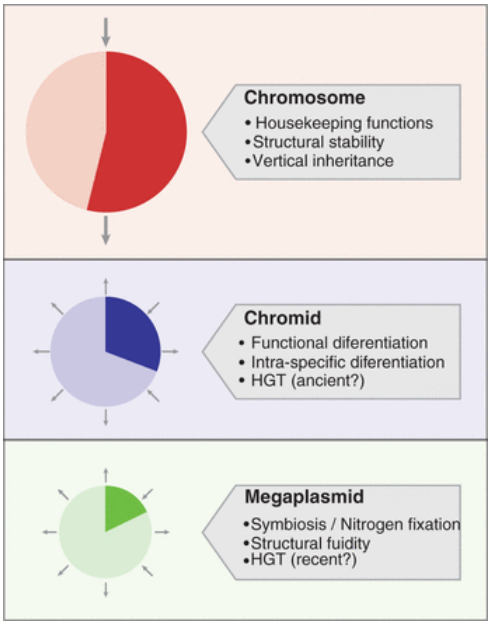


Figure 1.14 Tasks and evolutionary differences of replicons in *S. meliloti*. Pie charts indicate the proportion of each replicon that is present in the DNA backbones. The arrows indicate the transmission mechanism: vertical inheritance (two arrows) or HGT (radial arrows) from (Galardini et al., 2013b).

Recent experimental studies performed by diCenzo et al. (diCenzo et al., 2014) on strains obtained by the complete removal of pSymA, pSymB or of both of them, allowed to outline a more detailed evolutionary model of the multipartite genome in *Sinorhizobium meliloti* strains. In particular, the results indicated that the chromosome is sufficient for the growth in bulk soil environment, and suggested that pSymA and pSymB are more involved in metabolic capabilities for the growth in specialized ecological niche, as the rhizosphere and the symbiotic nodules.

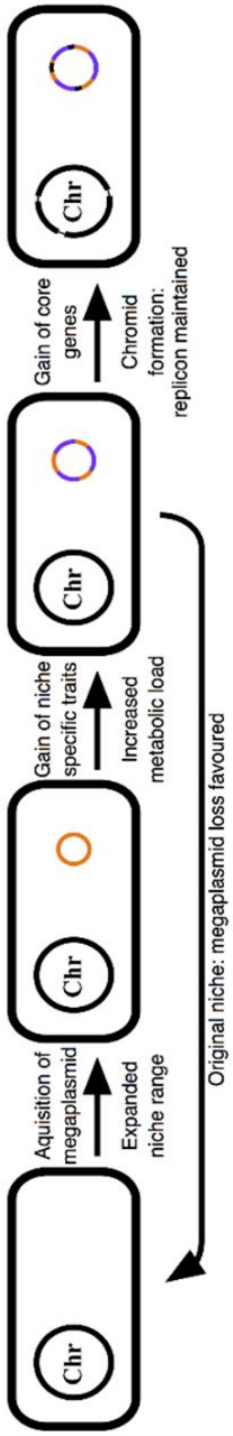


Figure 1.15 Schematic illustrating the described model of multipartite genome evolution and chromid formation from (diCenzo et al., 2014).

In the proposed model, the ancestral *S. meliloti* acquired a plasmid (the proto-pSymB megaplasmid), which encoded particular genetic traits allowing the cell to expand its ecological niche (e.g. to colonize the rhizosphere environment). Subsequently, the selective pressure of the new environment allowed the replicon to acquire additional genes through HGT providing a fitness benefit for this new location. This new large amount of genetic material acquired by the proto-megaplasmid, results in a metabolic advantage for the cell in the new environment, but contrarily, it could disadvantage the bacterium in its original niche. The translocation and integration of essential genes from the chromosome into the proto-megaplasmid resulted in the composition of a new element, the chromid, now maintained during evolution for the essential functions which holds, making it indispensable in all environments. Later on, this ancestral rhizospheric *S. meliloti* cell, acquired through HGT the proto-pSymA megaplasmid, which allowed ultimately to enter in symbiosis with the host legumes.

1.6 The puzzling dispensable genome of *S. meliloti*

From recent studies, it has become clear that the gene content of the bacterial dispensable genome is commonly responsible for the evolutionary diversification and the adaptive capability of different strains (Galardini et al., 2011;Galardini et al., 2013b). Despite the large interest on the composition of this genetic fraction (often larger than the core), the function and the role of some genetic determinants remain still not totally clarified, especially those in relation with the bacterial lifestyle and the phenotypic characteristics which characterize every particular strain. In *S. meliloti* several essential features of the symbiotic association with its host plant must be still elucidated. Recent studies evidenced that particular characteristics linked to environmental conditions or to stresses adaptation can be present in unrelated organisms. This could be the case of the heavy metal resistance, conferred by several types of genes, as for example *cnr*, *nre*, *czc* systems (Mergeay et al., 2003) (Grass et al., 2001).These determinants are present in heavy metal-resistant bacteria, but, interestingly, also in many non-resistant strains (Mengoni et al., 2010). Their function in those strains is still unclear, in particular in relation with the genomic localization. Indeed, *nreB* gene (encoding for a Ni+/H+ antiporter) in *S. meliloti* strains is located on pSymA plasmid, and its position leaves to hypostasize a possible involvement in the symbiotic capabilities of the strains, possibly in the metal balance for the symbiosome development (Pini et al., 2013).

Among the *S. meliloti* genes included in the dispensable genome and of which the function related with the symbiotic lifestyle is not totally elucidated, that encoding the enzyme ACC (1-

aminocyclopropane-1-carboxylate) deaminase is one of the most puzzling.

1.6.1 What does it mean having an ACC deaminase?

The plant response to stresses is mediated by its hormones, which can regulate the expression of specific proteins, necessary for the protection of plant cells from the deleterious effects they caused (Singh et al., 2015). As previously reported, PGPB can indirectly or directly interfere with the plant growth (Glick, 1995), and commonly, the bacterial production of specific compounds is able to modulate the stresses response in the plant. One of the most common plant hormone that mediates response to the stressors is ethylene. This is a gaseous hormone with a large range of biological effect on the plant in its developmental process, as the seeds germination, the tissues differentiation, the roots formation and elongation, the flowers blossoming and the fruits ripening (Abeles et al., 1992). Therefore, the plant senescence is clearly influenced by this hormone. A local increase of the concentration of ethylene can also occur during stress signalling and during the plant-bacteria symbiosis. Indeed, this interaction can benefit the host but, at the same time, it is an infection that can alter the electrochemical and biological balance of the plant.

Numerous studies suggested the importance of bacterial multimeric enzyme ACC (1-aminocyclopropane-1-carboxylate) deaminase in the reduction of the plant-ethylene level. Indeed, it was postulated (Glick et al., 1998) that the ACC deaminase-producing PGPB are able to bind the surface of plant roots or seeds, and once penetrated inside the cells, they can take up the free portion of ACC (amino

acid which is the precursor of ethylene), with the subsequent sequestering and cleaving of it. The enzyme ACC deaminase is encoded by the *acdS* gene, present in the dispensable genome fraction of most of the rhizobial strains. It was hypothesized that bacterial ACC deaminase is able to hydrolase ACC to ammonia and α-keto-butyrate decreasing the total amount of ACC in the plant. This would interrupt the normal cycle for the production of ethylene, reducing the extent of the hormone effect.

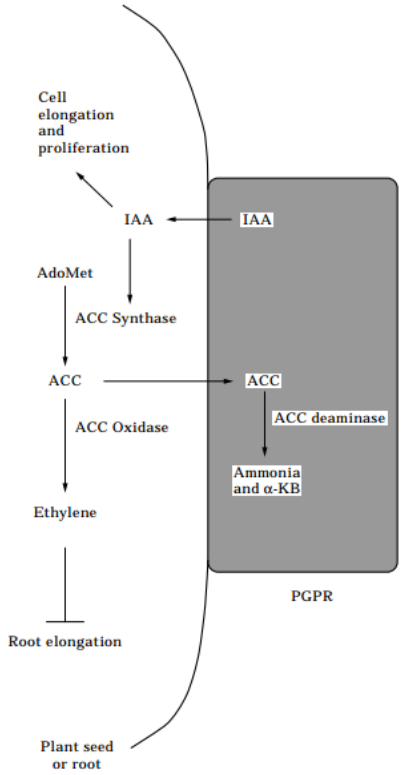


Figure 1.16 Schematic representation of how ACC deaminase producing PGPB can influence the level of plant ethylene from (Glick et al., 1998).

The presence of ACC deaminase was confirmed in a large set of bacterial species, and several models for its activity and regulation

were proposed (Li and Glick, 2001;Ma et al., 2003a;Duan et al., 2009;Duan et al., 2013) even if only a small fraction of putative *acdS* genes has been shown to encode active enzyme (Nascimento et al., 2013;Nascimento et al., 2014). Furthermore, the transport processes, which involves the sequestering of ACC and its displacement from the bacterial cells to the plant cells, have never been demonstrated in any plant-microbe symbiotic system. Furthermore, even if in several reports the positive effect of possessing ACC deaminase has been reported, its beneficial effect cannot be generalized to every bacterial species and to all nodulating plants (Murset et al., 2012).

1.7 Outline of the thesis

The nitrogen fixing symbiosis involves most of the 18,000 legume species together with an expanding collection of α and β proteobacteria, referred as rhizobia. This cooperation has a high ecological importance, occurs on all continents and accounts for a fourth of the global nitrogen fixed annually. The knowledge on the BNF operated by symbiotic bacteria has increased the interest on the application and the manipulation of them, in particular on the ecological and genomic evolutionary mechanisms which are the basis of the symbiotic association

The studies presented in this thesis aimed to give an insight into the bacteria lifestyle associated with host plant.

i) One of the main topics presented in this thesis is the sociomicrobiological interactions established by bacteria with other bacteria or plants. By now, it is clear that with rare exceptions microbes did not practice sociality (Greenberg, 2011). Just the terminology “colony growth” implies a sort of social interaction. Bacterial cells cooperate and communicate to perform a large range of multicellular behavior, such biofilm formation and quorum sensing (West et al., 2006). It is assumed that cooperation is favored during evolution because it can provide a benefit at population or a species level. Actually, this is not totally true. Several studies argued that the formation of a population can be a risk when single individuals (cheaters or free loaders) don’t cooperate but obtain benefits from the interaction with the other (Hamilton, 1964). In addition, cheaters can be competitive and be able to invade and take over the population, supporting the economical and human parallelism with the tragedy of the commons (Hardin, 2009). That’s

why explaining cooperation is one of the greatest problems for evolutionary biologists, which are mainly focused on the consequences of a social behavior and on the fitness improvement. Answer to the question: “how can cooperative behavior be maintained?” could help to find an explanation for the natural selection of the cooperative behavior.

In the rhizobia-legumes symbiosis, the interaction between bacteria and its host from both the ecological and the molecular point of views has largely been explored. The plant benefits from the nitrogen supplied by symbiotic bacteria, that is able to growth and use nutrients provided by the host. Several experiments have shown that this cooperation is favored, even if energetically expensive for both the partners. Indeed, if the rhizobia in a nodule doesn’t provide the host with fixed nitrogen, the plant punishes them by decreasing the O_2 in the nodule, which severely reduces the growth rate of the bacteria (Kiers et al., 2003). In this case the cheater (not able to fix N_2) has a direct fitness cost. This cost occurs at nodule level, that usually seems to be colonized by a single clone. Actually, recent studies explained that mixed nodules can occur, but this was not really demonstrated (Denison, 2000). The presence of mixed nodules could indeed explain the reduction of the evolutionary effects of nodule-level sanctions. When the cheater shares a nodule with a mutualist is somewhat protected from nodule-level sanctions.

In the chapter 2, we tried to give evidence of the mixed nodules existence and of the consequences of the cheater behavior inside the same root nodule. The study deepens the sociomicrobiological behavior of the free rider when it is in presence of mutualists strains.

ii) Another issue addressed in this thesis was related to deciphering the functions of dispensable genes located on the symbiotic megaplasmid in relation to the symbiosis. In particular, we focus on the expression and regulation of the gene *acdS*, located in the dispensable genome of a large variety of rhizobial strains. This gene, able to encode the enzyme 1-aminocyclopropane-1-carboxylate (ACC) deaminase, seems to be linked to the improvement of symbiotic performance and competitiveness of bacterial strains (Ma et al., 2003b;Ma et al., 2004). Indeed, the enzyme is responsible for the sequestering and cleaving of ACC, the immediate precursor of the hormone ethylene. The systemic lowering of the ethylene production allows a decrease of its inhibitory effect on the plant growth and on the root elongation. Up to now, the activity of the enzyme and the presence of *acdS* gene were tested in *Pseudomonas* sp. strain ACP and in a wide range of microbes including the fungus *Penicillium citrinum*, the yeast *Hansenula saturnus* (Honma and Shimomura, 1978). Furthermore, the enzyme was found in a large number of plant growth promoting rhizobacteria, including *Rhizobium leguminosarum* bv. *viciae*, *Rhizobium hedysari*, and *Mesorhizobium loti* (Duan et al., 2009). For *Sinorhizobium meliloti*, past works (Ma et al., 2004) showed the increase of the nodule competitiveness of a recombinant strain containing *acdS* gene respect to the wild type, or the increase of the host plant growth when infected by ACC deaminase-overproducing *S. meliloti* strains (Kong et al., 2015). However, no details on the functional role and regulation of the gene in *S. meliloti* native strains are present, especially in relation with their symbiotic performance.

In the chapter 3, we tried to explore the phylogenetic profile of *acdS* orthologs among rhizobia strains and the functional role of the gene in relation with its symbiotic behavior.

iii) While most microbial genomes have only a unique chromosome, many are more complex and consist of two or more large replicons. This organization is largely diffused among proteobacteria, but it is not limited to this class. Indeed, a complex multipartite genome is present in plant symbionts as *Sinorhizobium* and *Rhizobium* species, and plant and animal pathogens (*Agrobacterium*, *Vibrio*, *Burkholderia* and *Brucella*). Usually, the chromosome contains all the genes essential for the free-living growth of the organism, indeed classified as core genes. Instead, the "non-essential genes", involved in specific metabolic function linked to stresses response or adaptive traits, are usually located on the secondary replicons. In *Sinorhizobium meliloti* strains, the secondary replicons which are part the genome are named megaplasmid (laterally acquired with a plasmid origin of replication and characterized by the complete absence of core genes), and chromid, which shows characteristics of both chromosomes and megaplasmids (Harrison et al., 2010).

To understand the role, the evolution and the metabolic connections of these replicons could provide an important deepening in the biology of the bacterium, in particular in relation with its symbiotic lifestyle and all the possible strategies to promote or modified the interactions.

Recently, a possible scenario for the evolution of *S. meliloti* multipartite genome was proposed (diCenzo et al., 2014) and several hypothesis were advanced for the function of every replicons, in particular related with the environmental specialization

and microbial life cycle. Indeed, the bacterial genome are not randomly organized. However, these proposals were not really demonstrated.

In chapter 4, we tried to give an explanation and a confirmation of the niche specialization hypothesis through a genome scale metabolic modelling. In particular, we focus on the metabolic role and the fitness contribute of genes carried by each replicon during the shifts between the three environment colonisable by *S. meliloti*: bulk soil, rhizosphere and symbiotic nodule.

iv) In this thesis the appendix is reporting additional works related with the plant-microbe's interaction. In particular, the studies focused on:

- On the bacterial microbiota (essentially endophytes) associated with aromatic plants from the genus *Thymus*. In this work, we tried to answer to the questions on the possible association between the endophytic microbial community and the essential oil extracted from their respective host plant. In particular, we focused on the pattern of resistance of the plants isolates to the EOs, considering their species specific chemical composition.
- We prepared a literature review on the use of rhizobial symbiont to improve phytoremediation in contaminated areas.
- We contribute to a book chapter describing the pipeline of genomic comparative analysis by the use of IMG tool.

Bibliography

Abeles, Morgan, and Saltveit (1992). "Ethylene in Plant Biology," ed. E. 2. (Academic Press, New York).

Akçay, and Simms (2011). Negotiation, sanctions, and context dependency in the legume-rhizobium mutualism. *The American Naturalist* 178(1)**,** 1-14.

Allen, and Allen (1981). The Leguminosae: A source book of characteristics , uses and nodulation. *University of Wisconsin Press, Madion, WJ/McMillian Publishing*.

Alvin, A., Miller, K.I., and Neilan, B.A. (2014). Exploring the potential of endophytes from medicinal plants as sources of antimycobacterial compounds. *Microbiological Research* 169**,** 483-495. doi: 10.1016/j.micres.2013.12.009.

Atkins, a.S. (2000). "Ureide synthesis in legume nodules.," in *Prokaryotic Nitrogen Fixation,* ed. Triplett. Wymondham: Horizon Scientific).

Bacon, and Hinton (2006). "Bacterial Endophytes: The Endophytic Niche, its occupants, and its utility," in *Plant-Associated Bacteria,* ed. S.S.E. Gnanamanickam. (The Netherland Springer), 155-194.

Barnett, and Fisher (2006). Global gene expression in the rhizobial-legume symbiosis. *Symbiosis (Rehovot)* 42(1)**,** 1-24.

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

Bashan, and Holguin (1998). Proposal for the division of plant growth-promoting rhizobacteria into two classifications: biocontrol-PGPB (plant growth-promoting bacteria) and PGPB. *Soil Biology and Biochemistry***,** 1225-1228.

Batut, Terzaghi, Gherardi, Huguet, Terzaghi, Garnerone, and Huguet (1985). Localization of a symbiotic fix region on Rhizobium meliloti pSym megaplasmid more than 200

kilobases from the nod-nif region. *Molecular and General Genetics* 232-239.

Becker (2007). "Functional genomics of rhizobia," in *Prokaryotic Symbionts in Plants*. Springer Berlin Heidelberg), (pp. 71-100).

Berg, and Hallmann (2006). "Control of plant pathogenic fungi with bacterial endophytes," in *Microbial root endophytes,*, ed. C.B. B. Schulz, T. Sieber (Eds.). Springer-Verlag, Berlin), pp. 53-69.

Berg, G., et al. (2005). Endophytic and ectophytic potato-associated bacterial communities differ in structure and antagonistic function against plant pathogenic fungi. *FEMS Microbiology Ecology* 51**,** 215-229.

Berleman, Scott, Chumley, and Kirby (2008). Predataxis behavior in Myxococcus xanthus. . *P Natl Acad Sci USA* 105**,** 17127-17132.

Bouton, J. (2007). The economic benefits of forage improvement in the United States. *Euphytica* 154**,** 263-270. doi: 10.1007/s10681-006-9220-6.

Bowen , Beerling DJ, Koch PL, Z.J., and T., Q. (2004). A humid climate state during the Palaeocene/Eocene thermal maximum. *Nature* 432(7016)**,** 495-499.

Brewin (2004). Plant cell wall remodelling in the Rhizobium-legume symbiosis. *Critical Reviews in Plant Sciences* 23(4)**,** 293-316.

Brewin, N.J. (1991). Development of the legume root nodule. *Annu Rev Cell Biol* 7**,** 191-226. doi: 10.1146/annurev.cb.07.110191.001203.

Capela, D., Filipe, C., Bobik, C., Batut, J., and Bruand, C. (2006). *Sinorhizobium meliloti* differentiation during symbiosis with alfalfa: a transcriptomic dissection. *Mol Plant Microbe Interact* 19**,** 363-372. doi: 10.1094/MPMI-19-0363.

Chen, Moulin, Bontemps, Vandamme, Béna, and Boivin-Masson (2003). Legume symbiotic nitrogen fixation byβ-

proteobacteria is widespread in nature. *Journal of bacteriology* 185(24)**,** 7266-7272.

Christina, Christapher, and Bhore (2013). Endophytic bacteria as a source of novel antibiotics: an overview. *Pharmacognosy reviews,* 7(13)**,** 11.

Cosson, and Soldati (2008). Eat, kill or die: when amoeba meets bacteria. . *Current opinion in microbiology* **,** 271-276.

Deakin, and Broughton (2009). Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nature Reviews Microbiology* 7(4), 312-320.

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

diCenzo, MacLean, Milunovic, Golding, and Finan (2014). Examination of prokaryotic multipartite genome evolution through experimental genome reduction. *PLoS Genet* 10(10).

Dobert, R.C., Breil, B. T., & Triplett, E. W. (1994). DNA sequence of the common nodulation genes of Bradyrhizobium elkanii and their phylogenetic relationship to those of other nodulating bacteria. *MPMI-Molecular Plant Microbe Interactions* 7(5)**,** 564-572.

Dresler-Nurmi, Fewer, Räsänen, and Lindström (2007). "The diversity and evolution of rhizobia," in *Prokaryotic Symbionts in Plants*. (Springer Berlin Heidelberg), (pp. 3-41).

Duan, Jiang, Cheng, Heikkila, and Glick (2013). The complete genome sequence of the plant growth-promoting bacterium Pseudomonas sp. UW4. . *PLoS ONE* 8:e58640.

Duan, Müller, Charles, Vesely, and Glick (2009). 1-Aminocyclopropane-1-Carboxylate (ACC) Deaminase Genes in Rhizobia from Southern Saskatchewan. *Microbial Ecology* 57**,** 423-436. doi: 10.1007/s00248-008-9407-6.

Fischer (1994). Genetic regulation of nitrogen fixation in rhizobia. *Microbiological reviews* 58(3)**,** 352-386.

Galardini, Bazzicalupo, Biondi, Brambilla, Brilli, Bruce, Chain, Chen, Daligault, Walston Davenport, Deshpande, Detter, J., Goodwin, Han, Han, Huntemann, Ivanova, Klenk, Kyrpides, Markowitz, Mavrommatis, Mocali, Nolan, Pagani, Pati, Pini, Pitluck, Spini, Szeto, Teshima, Woyke, and Mengoni (2013a). Permanent draft genome sequences of the symbiotic nitrogen fixing Ensifer meliloti strains BO21CC and AK58. *Standards in Genomic Sciences* 9**,** 2.

Galardini, Mengoni, Brilli, Pini, Fioravanti, Lucas, Lapidus, Cheng, Goodwin, Pitluck, Land, Hauser, Woyke, Mikhailova, Ivanova, Daligault, Bruce, Detter, Tapia, Han, Teshima, Mocali, Bazzicalupo, and Biondi (2011). Exploring the symbiotic pangenome of the nitrogen-fixing bacterium *Sinorhizobium meliloti*. *BMC Genomics* 12**,** 235.

Galardini, Pini, Bazzicalupo, Biondi, and Mengoni (2013b). Replicon-Dependent Bacterial Genome Evolution: The Case of *Sinorhizobium meliloti*. *Genome Biology and Evolution* 5**,** 542-558. doi: 10.1093/gbe/evt027.

Galibert, F., Finan, T.M., Long, S.R., Puhler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M.J., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R.W., Dreano, S., Federspiel, N.A., Fisher, R.F., Gloux, S., Godrie, T., Goffeau, A., Golding, B., Gouzy, J., Gurjal, M., Hernandez-Lucas, I., Hong, A., Huizar, L., Hyman, R.W., Jones, T., Kahn, D., Kahn, M.L., Kalman, S., Keating, D.H., Kiss, E., Komp, C., Lelaure, V., Masuy, D., Palm, C., Peck, M.C., Pohl, T.M., Portetelle, D., Purnelle, B., Ramsperger, U., Surzycki, R., Thebault, P., Vandenbol, M., Vorholter, F.J., Weidner, S., Wells, D.H., Wong, K., Yeh, K.C., and Batut, J. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* 293**,** 668-672. doi: 10.1126/science.1060966293/5530/668 [pii].

Gamalero, and Glick (2015). Bacterial modulation of plant ethylene levels. *Plant physiology* 169**,** 13-22.

Gibson, K.E., Kobayashi, H., and Walker, G.C. (2008). Molecular Determinants of a Symbiotic Chronic Infection. *Annual Review of Genetics* 42**,** 413-441.

Glick (1995). The enhancement of plant growth by free-living bacteria. *Canadian Journal of Microbiology* **,** 109-117.

Glick (2012). Plant growth-promoting bacteria: mechanisms and applications. *Scientifica*.

Glick, Penrose, and Li (1998). A model for the lowering of plant ethylene concentrations by plant growth-promoting bacteria. *Journal of Theoretical Biology* 190**,** 63-68.

Gogarten, J.P., Doolittle, W.F., and Lawrence, J.G. (2002). Prokaryotic evolution in light of gene transfer. *Mol Biol Evol* 19**,** 2226 - 2238.

Grass, G., Fan, B., Rosen, B.P., Lemke, K., Schlegel, H.G., and Rensing, C. (2001). NreB from *Achromobacter xylosoxidans* 31A Is a nickel-induced transporter conferring nickel resistance. *J Bacteriol* 183**,** 2803-2807. doi: 10.1128/JB.183.9.2803-2807.2001 [doi].

Greenberg, E. (2011). THE NEW SCIENCE OF SOCIOMICROBIOLOGY AND THE REALM OF SYNTHETIC AND SYSTEMS ECOLOGY.

Hamilton, W. (1964). The genetical evolution of social behaviour. *Journal of theoretical biology* 7(1)**,** 17-52.

Hardin, G. (2009). The tragedy of the commons. *Journal of Natural Resources Policy Research* 1(3)**,** 243-253.

Harrison, P.W., Lower, R.P.J., Kim, N.K.D., and Young, J.P.W. (2010). Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends in Microbiology* 18**,** 141-148. doi: 10.1016/j.tim.2009.12.010.

Honma, and Shimomura (1978). Metabolism of 1-aminocyclopropane-1-carboxylic acid. *Agric. Biol. Chem* 42**,** 1825-1831. doi: doi:10.1271/bbb1961.42.1825.

Jones, K.M., Kobayashi, H., Davies, B.W., Taga, M.E., and Walker, G.C. (2007). How rhizobial symbionts invade plants: the Sinorhizobium-Medicago model. *Nat Rev Microbiol* 5**,** 619-633. doi: nrmicro1705 [pii]10.1038/nrmicro1705.

Kiers, Palmer, Ives, Bruno, and Bronstein (2010). Mutualisms in a changing world: an evolutionary perspective. *Ecology Letters*, 13(12)**,** 1459-1474.

Kiers, E.T., Rousseau, R.A., and Denison, R.F. (2006). Measured sanctions: legume hosts detect quantitative variation in rhizobium cooperation and punish accordingly. *Evolutionary Ecology Research* 8**,** 1077-1086.

Kiers, E.T., Rousseau, R.A., West, S.A., and Denison, R.F. (2003). Host sanctions and the legume-rhizobium mutualism. *Nature* 425**,** 78-81. doi: 10.1038/nature01931.

Kneip, Lockhart, Voss, and Maier (2007). Nitrogen fixation in eukaryotes - new models for symbiosis. *BMC Evol. Biol.* 7, 55.

Kondorosi, Banfalvi, and Kondorosi (1984). Physical and genetic analysis of a symbiotic region of Rhizobium meliloti: identification of nodulation genes. *Molecular and General Genetics MGG* 193(3) **,** 445-452.

Kong, Mohamad, Deng, Liu, Glick, and Wei (2015). Rhizobial symbiosis effect on the growth, metal uptake, and antioxidant responses of Medicago lupulina under copper stress. *Environmental Science and Pollution Research* 22(16)**,** 12479-12489.

Koonin, E.V. (2009). Darwinian evolution in the light of genomics. *Nucl. Acids Res.* 37**,** 1011-1034. doi: 10.1093/nar/gkp089.

Laranjo, Alexandre, and Oliveira (2014). Legume growth-promoting rhizobia: an overview on the Mesorhizobium genus. *Microbiological research* 169(1)**,** 2-17.

Lavin M., Herendeen P. S., and F, W.M. (2005). Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Systematic biology* 54(4)**,** 575-594.

Li, and Glick (2001). Transcriptional regulation of the Enterobactercloaceae UW4 1-aminocyclopropane-1-carboxylate (ACC) deaminase gene (AcdS). *Canadian Journal of Microbiology* 47**,** 259-267.

Lodwig, and Poole (2003). Metabolism of Rhizobium bacteroids. *Critical reviews in plant sciences* 22(1)**,** 37-78.

Ma, Guinel, and Glick (2003a). Rhizobium leguminosarum biovarviciae 1- aminocyclopropane-1-carboxylate deaminase promotes nodulation of pea plants. *Appl. Environ. Microbiol.* 69**,** 4396-4402.

Ma, Sebestianova, Sebestian, Burd, Guinel, and Glick (2003b). Prevalence of 1-aminocyclopropane-1-carboxylate deaminase in Rhizobium spp. *Antonie van Leeuwenhoek* 83**,** 285-291.

Ma, Wenbo, Trevor, and Glick. (2004). Expression of an exogenous 1-aminocyclopropane-1-carboxylate deaminase gene in Sinorhizobium meliloti increases its ability to nodulate alfalfa. *Applied and Environmental Microbiology* 70**,** 5891-5897.

McCaslin, and Miller (2007). The future of alfalfa as a biofuels feedstock. . *In 37th California Alfalfa & Forage Symposium* (pp. 17-19).

Medini, D., Donati, C., Tettelin, H., Masignani, V., and Rappuoli, R. (2005). The microbial pan-genome. *Current Opinion in Genetics & Development* 15**,** 589-594. doi: 10.1016/j.gde.2005.09.006.

Mengoni, A., Schat, H., and Vangronsveld, J. (2010). Plants as extreme environments? Ni-resistant bacteria and Ni-hyperaccumulators of serpentine flora. *Plant and Soil* 331**,** 5-16.

Mergeay, M., Monchy, S., Vallaeys, T., Auquier, V., Benotmane, A., Bertin, F., Taghavi, S., Dunn, J., van der Lelie, D., and Wattiez, R. (2003). *Ralstonia metallidurans*, a bacterium specifically adapted to toxic metals: towards a catalogue of metal-responsive genes. *Fems Microbiology Reviews* 27**,** 385-410.

Moulin, L., Munive, A., Dreyfus, B., and Boivin-Masson, C. (2001). Nodulation of legumes by members of the beta-subclass of Proteobacteria. *Nature* 411**,** 948-950. doi: 10.1038/3508207035082070 [pii].

Moya, Peretó, Gil, and Latorre (2008). Learning how to live together: genomic insights into prokaryote-animal symbioses. . *Nature Reviews Genetics*, 9(3)**,** 218-229.

Mozaffari, M., Rosen, C.J., Russelle, M.P., and Nater, E.A. (2000). Corn and Soil Response To Application of Ash Generated From Gasified Alfalfa Stems. *Soil Science* 165**,** 896-907.

Murset, Hennecke, and Pessi (2012). Disparate role of rhizobial ACC deaminase in root-nodule symbioses. *Symbiosis* 57**,** 43-50.

Nascimento, Rossi, Soares, McConkey, and Glick (2014). New Insights into 1-Aminocyclopropane-1-Carboxylate (ACC) Deaminase Phylogeny, Evolution and Ecological Significance. *PLoS ONE* 9**,** e99168.

Nascimento, Vicente, Barbosa, Espada, Glick, Mota, and Oliveira (2013). Evidence for the involvement of ACC deaminase from Pseudomonas putida UW4 in the biocontrol of pine wilt disease caused by Bursaphelenchus xylophilus. . *BioControl*, 58(3) **,** 427-433.

O'Brian, Kirshbom, and Maier (1987). Bacterial heme synthesis is required for expression of the leghemoglobin holoprotein but not the apoprotein in soybean root nodules. *Proceedings of the National Academy of Sciences***,** 8390-8393.

Partida-Martinez, and Heil (2011). The microbe-free plant: fact or artifact? *Frontiers in plant science* 2**,** 100.

Peters, Fisher, and Dean (1995). Nitrogenase structure and function: a biochemical-genetic perspective. *Annual Reviews in Microbiology***,** 335-366.

Peters, and Szilagyi (2006). Exploring new frontiers of nitrogenase structure and mechanism. *Current opinion in chemical biology***,** 101-108.

Pini, Galardini, Bazzicalupo, and Mengoni (2011). Plant-bacteria association and symbiosis: are there common genomic traits in *Alphaproteobacteria*? *Genes* 2**,** 1017-1032.

Pini, Spini, Galardini, Bazzicalupo, Benedetti, Chiancianesi, Florio, Lagomarsino, Migliore, Mocali, and Mengoni (2013). Molecular phylogeny of the nickel-resistance gene *nreB* and functional role in the nickel sensitive symbiotic nitrogen fixing bacterium *Sinorhizobium meliloti*. *Plant and Soil* doi: 10.1007/s11104-013-1979-3.

Pladys, and Vance (1993). Proteolysis during Development and Senescence of Effective and Plant Gene-Controlled Ineffective Alfalfa Nodules. *Plant Physiol. Biochem* 103(2):379-384.

Postgate (1982). The fundamentals of nitrogen fixation. *CUP Archive*.

Potts, Blesmeijer, Kremen, Neumann, Schweiger, and Kunin (2010). Global pollinator declines: trends, impacts and drivers. *Trends Ecol. Evol.* 25**,** 345-353.

Provorov, N.A. (1998). Coevolution of rhizobia with legumes: facts and hypotheses. *Symbiosis, Philadelphia, Pa.(USA)*.

Puppo, Groten, Bastian, Carzaniga, Soussi, Lucas, and Foyer (2005). Legume nodule senescence: roles for redox and hormone signalling in the orchestration of the natural aging process. . *New Phytologist*, 165(3)**,** 683-701.

Raymond, Siefert, Staples, and Blankenship (2004). The natural history of nitrogen fixation. *Molecular biology and evolution* 21(3)**,** 541-554.

Reed, and Glick (2013). Applications of plant growth-promoting bacteria for plant and soil systems. *Applications of Microbial Engineering* Taylor and Francis, Enfield, CT**,** 181-229.

Reinhold-Hurek, B., and Hurek, T. (2011). Living inside plants: bacterial endophytes. *Current Opinion in Plant Biology* 14**,** 435-443. doi: http://dx.doi.org/10.1016/j.pbi.2011.04.004.

Rodríguez-Haas, Finney, Vogt, González-Melendi, Imperial, and González-Guerrero (2013). Iron distribution through the developmental stages of Medicago truncatula nodules. . *Metallomics* 5(9)**,** 1247-1253.

Roponen (1970). The effect of darkness on the leghemoglobin content and amino acid levels in the root nodules of pea plants. *Physiol Plant* 23**,** 452-460.

Rosenblueth, M., and Martinez-Romero, E. (2006). Bacterial endophytes and their interactions with hosts. *Mol Plant Microbe Interact* 19**,** 827-837. doi: 10.1094/MPMI-19-0827.

Roumiantseva, Andronov, Sharypova, Dammann-Kalinowski, Keller, Young, and Simarov (2002). Diversity of Sinorhizobium meliloti from the Central Asian Alfalfa Gene Center. *Applied and Environmental Microbiology***,** 4694-4697.

Schneiker-Bekel, Wibberg, Bekel, Blom, Linke, Neuweger, Stiens, Vorhölter, Weidner, Goesmann, Pühler, and Schlüter (2011). The complete genome sequence of the dominant *Sinorhizobium meliloti* field isolate SM11 extends the *S. meliloti* pan-genome. *Journal of Biotechnology* 155**,** 20-33. doi: doi:10.1016/j.jbiotec.2010.12.018.

Schrire B.D., Lewis G., and M., L. (2005). "Biogeography of the Leguminosae," in *Legumes of the world* ed. S.B. Lewis G, Mackinder B, Lock M. (London: Royal Botanic Gardens), 21-54.

Serova, and Tsyganov (2014). Symbiotic nodule senescence in legumes: molecular-genetic and cellular aspects (review) *Сельскохозяйственная биология* 5.

Shaw, Morris, and Hooker (2006). Perception and modification of plant flavonoid signals by rhizosphere microorganisms. *Environmental Microbiology* 8(11)**,** 1867-1880.

Singh, Shelke, Kumar, and Jha (2015). Biochemistry and genetics of ACC deaminase: a weapon to “stress ethylene” produced in plants. *Frontiers in microbiology* 6.

Sprent (1980). Root nodule anatomy, type of export product and evolutionary origin of some Leguminosae. *Plant , Cell and Environment* 3**,** 35-43.

Sprent, and Gehlot (2010). Nodulated legumes in arid and semi-arid environments: are they important? *Plant Ecology & Diversity***,** 211-219.

Sprent, J.I. (1994). Evolution and diversity in the legume-rhizobium symbiosis: chaos theory? *Plant and Soil* 161**,** 1-10.

Sprent, J.I. (2001). *Nodulation in legumes.* London: Royal Botanic Gardens, Kew.

Sprent, J.I. (2007a). Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. *New Phytol* 174**,** 11-25.

Sprent, J.I. (2007b). Evolving ideas of legume evolution and diversity: A taxonomic perspective on the occurrence of nodulation: Tansley review. *New Phytologist* 174**,** 11-25.

Sprent, J.I. (2008). 60Ma of legume nodulation: What's new? What's changing? *J Exp Bot* 59**,** 1081-1084. doi: erm286 [pii]10.1093/jxb/erm286.

Sprent, J.I., and James, E.K. (2007). Legume evolution: where do nodules and mycorrhizas fit in? *Plant Physiol* 144**,** 575-581. doi: 144/2/575 [pii]10.1104/pp.107.096156.

Stewart (1966). Nitrogen fixation in plants.

Stingl, Radek, Yang, and Brune (2005). Endomicrobia: cytoplasmic symbionts of termite gut protozoa form a separate phylum of prokaryotes. *Appl. Environ. Microbiol.* 71**,** 1473-1479

Stracke, Kistner, Yoshida, Mulder, Sato, Kaneko, and Parniske (2002). A plant receptor-like kinase required for both bacterial and fungal symbiosis. . *Nature,* 417(6892)**,** 959-962.

Streng, op den Camp, Bisseling, and Geurts (2011). Evolutionary origin of rhizobium Nod factor signaling. . *Plant signaling & behavior* 6(10)**,,** 1510-1514.

Sturz, and Nowak (2000). Endophytic communities of rhizobacteria and the strategies required to create yield enhancing associations with crops. *Applied Soil Ecology* 15**,** 183-190.

Taylor, Radax, Steger, and Wagner (2007). Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. *Microbiol.Mol. Biol. Rev.***,** 295-347.

Tettelin, H., Riley, D., Cattuto, C., and Medini, D. (2008). Comparative genomics: the bacterial pan-genome. *Current Opinion in Microbiology* 11**,** 472-477.

Trainer, and Charles (2006). The role of PHB metabolism in the symbiosis of rhizobia with legumes. . *Applied microbiology and biotechnology,* 71(4) **,** 377-386.

Turner, and Young (2000). The glutamine synthetases of rhizobia: phylogenetics and evolutionary implications. . *Molecular Biology and Evolution,* 17(2)**,** 309-319.

Van de Velde, Guerra, De Keyser, De Rycke, Rombauts, Maunoury, and Goormachtig (2006). Aging in legume symbiosis. A molecular view on nodule senescence in Medicago truncatula. . *Plant Physiology* 141(2)**,** 711-720.

van Loon, L.C., Bakker, P.A.H.M., and Pieterse, C.M.J. (1998). SYSTEMIC RESISTANCE INDUCED BY RHIZOSPHERE BACTERIA. *Annual Review of Phytopathology* 36**,** 453-483. doi: doi:10.1146/annurev.phyto.36.1.453.

Vance (2000). "Amide biosynthesis in root nodules of temperate legumes," in *Prokaryotic nitrogen fixation,* ed. Triplett. Wymondham: Horizon Scientific Press), pp. 589-607.

Werner, Cornwell, W.K., Sprent J. I., Kattge, J., and Kiers, E.T. (2014a). A single evolutionary innovation drives the deep evolution of symbiotic N2-fixation in angiosperms. . *Nature communications,* 5.

Werner, G.D.A., Strassmann, J.E., Ivens, A.B.F., Engelmoer, D.J.P., Verbruggen, E., Queller, D.C., Noe, R., Johnson, N.C., Hammerstein, P., and Kiers, E.T. (2014b). Evolution of microbial markets. *Proceedings of the National Academy of Sciences* 111**,** 1237-1244. doi: 10.1073/pnas.1315980111.

West, S., Griffin, A., Gardner, A., and Diggle, S. (2006). Social evolution theory for microorganisms. *Nature Reviews Microbiology***,** 597-607.

Wilson (2000). *Sociobiology.* Harvard University Press.

Wilson, Rice, Rillig, Springer, and Hartnett (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.

Wong, and Golding (2003). A phylogenetic analysis of the pSymB replicon from the Sinorhizobium meliloti genome reveals a complex evolutionary history. *Canadian Journal of Microbiology***,** 269-280.

Young, J.P., Crossman, L., Johnston, A., Thomson, N., Ghazoui, Z., Hull, K., Wexler, M., Curson, A., Todd, J., Poole, P., Mauchline, T., East, A., Quail, M., Churcher, C., Arrowsmith, C., Cherevach, I., Chillingworth, T., Clarke, K., Cronin, A., Davis, P., Fraser, A., Hance, Z., Hauser, H., Jagels, K., Moule, S., Mungall, K., Norbertczak, H., Rabbinowitsch, E., Sanders, M., Simmonds, M., Whitehead, S., and Parkhill, J. (2006). The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biology* 7**,** R34.

Chapter 2

2.1 Mixed nodule infection in *Sinorhizobium meliloti- Medicago sativa* symbiosis suggest the presence of cheating behaviour

Last decades have seen an increasing interest on the ecological and molecular aspects of plant-bacteria interactions. In particular, one of the main topic of the last few years has been related to the sociobiological issues arisen from such interaction in terms of antagonism or cooperation between bacterial strains and the plant. In the nitrogen fixing symbiosis between rhizobia and legumes, the mutualism taking place between plant and bacterium, as well as the exclusiveness of the interaction, is a key point for a successful symbiosis. Until now, the possibility to find different rhizobial strains in the same root nodule has not been experimentally verified. The occurrence of mixed nodules could represent an interesting chance for strains with different symbiotic performance to overcome the sanctions imposed by plants to the unproductive nodules. Finally, this strategy could allow to the non-mutualist strains to persist during the evolution.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 13 June 2016
doi: 10.3389/fpls.2016.00835

Mixed Nodule Infection in *Sinorhizobium meliloti–Medicago sativa* Symbiosis Suggest the Presence of Cheating Behavior

*Alice Checcucci[1], Elisa Azzarello[2], Marco Bazzicalupo[1], Marco Galardini[3], Alessandra Lagomarsino[4], Stefano Mancuso[2], Lucia Marti[2], Maria C. Marzano[2], Stefano Mocali[4], Andrea Squartini[5], Marina Zanardo[5] and Alessio Mengoni[1]**

[1] Department of Biology, University of Florence, Florence, Italy, [2] Department of Agri-Food Production and Environmental Science, University of Florence, Florence, Italy, [3] European Molecular Biology Laboratory – European Bioinformatics Institute, Wellcome Trust Genome Campus, Cambridge, UK, [4] Consiglio per la Ricerca in Agricoltura e l'Analisi dell'Economia Agraria, Centro di Ricerca per l'Agrobiologia e la Pedologia, Florence, Italy, [5] Department of Agronomy, Food, Natural Resources, Animals and the Environment, University of Padua, Padova, Italy

OPEN ACCESS

Edited by:
Stéphane Hacquard, Max Planck Institute for Plant Breeding Research, Germany

Reviewed by:
Simona Radutoiu, Aarhus University, Denmark
Pablo Vinuesa, Centro de Ciencias Genomicas – UNAM, Mexico

***Correspondence:**
Alessio Mengoni
alessio.mengoni@unifi.it

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

Received: *29 February 2016*
Accepted: *27 May 2016*
Published: *13 June 2016*

Citation:
Checcucci A, Azzarello E, Bazzicalupo M, Galardini M, Lagomarsino A, Mancuso S, Marti L, Marzano MC, Mocali S, Squartini A, Zanardo M and Mengoni A (2016) Mixed Nodule Infection in Sinorhizobium meliloti-Medicago sativa Symbiosis Suggest the Presence of Cheating Behavior. Front. Plant Sci. 7:835. doi: 10.3389/fpls.2016.00835

In the symbiosis between rhizobia and legumes, host plants can form symbiotic root nodules with multiple rhizobial strains, potentially showing different symbiotic performances in nitrogen fixation. Here, we investigated the presence of mixed nodules, containing rhizobia with different degrees of mutualisms, and evaluate their relative fitness in the *Sinorhizobium meliloti–Medicago sativa* model symbiosis. We used three *S. meliloti* strains, the mutualist strains Rm1021 and BL225C and the non-mutualist AK83. We performed competition experiments involving both *in vitro* and *in vivo* symbiotic assays with *M. sativa* host plants. We show the occurrence of a high number (from 27 to 100%) of mixed nodules with no negative effect on both nitrogen fixation and plant growth. The estimation of the relative fitness as non-mutualist/mutualist ratios in single nodules shows that in some nodules the non-mutualist strain efficiently colonized root nodules along with the mutualist ones. In conclusion, we can support the hypothesis that in *S. meliloti–M. sativa* symbiosis mixed nodules are formed and allow non-mutualist or less-mutualist bacterial partners to be less or not sanctioned by the host plant, hence allowing a potential form of cheating behavior to be present in the nitrogen fixing symbiosis.

Keywords: ***Medicago sativa*, symbiotic nitrogen fixation, cheating, mixed nodules, competition, *Sinorhizobium meliloti***

INTRODUCTION

In the last decades, the interest on social interaction strategies of bacteria, including antagonism and cooperation has increased (Kiers et al., 2006). Nitrogen fixing symbioses between rhizobia and leguminous plants provide interesting models to study social dynamics of strains, which compete for entering in symbiosis with the same host plant (Denison and Kiers, 2004). Rhizobia interact with plant roots after the perception of flavonoid molecules released by the plant roots. In response to flavonoid, lipo-chito-oligosaccharide molecules (Nod Factors) are produced by rhizobia and trigger a molecular pathway on plant cells. This mechanism ultimately leads to rhizobial entry

into plant root tissues, intracellular colonization, formation of the root nodule structure, and differentiation of intracellular rhizobia in bacteroids. Then, bacteroids will express the nitrogenase genes responsible for the fixation of atmospheric di-nitrogen to ammonium, thus providing a selective advantage to plants growing in nitrogen-depleted soils.

The plant–rhizobia relationship is not exclusive: multiple rhizobial strains colonize the same individual plant (Kiers et al., 2006), and large genetic differences in strains isolated from the very same root apparatus have been reported (Carelli et al., 2000). Then, in theory, plants can enter into symbiosis with rhizobial strains having different symbiotic performances (e.g., capacity to fix di-nitrogen, to colonize root nodules, etc.). To control the efficiency of the symbiotic partnership and then avoid overcolonization by strains fixing low amounts of nitrogen, plants have evolved mechanisms which favor nodules colonized by rhizobia supplying the host with a higher amount of nitrogen (Kiers et al., 2003). In fact, the plant seems to monitor bacterial performances and states sanctions on nitrogen fixation defective strains, penalizing them in the colonization (Kiers et al., 2003). However, in nature, strains belonging to the same species but with different levels of symbiotic performances are present. Such differences have also a practical importance in agronomy, since they are the basis for the selection of 'élite' inoculant strains to be applied for legume yield improvement (Dwivedi et al., 2015). Deciphering the competition patterns among natural strains and the selection by host plant would be of great importance for inoculant strains development.

An additional level of competition occurs at single nodule level. Several studies have reported the occurrence of a bacterial community composed by symbiotic rhizobia and by other (apparently non-symbiotic) strains in soybean, common bean, cowpea, and clover (Moawad and Schmidt, 1987; Mårtensson et al., 1989; Sessitsch et al., 1997; Denison, 2000). The colonization of the same nodule by different symbiotic strains may allow to establish a sort of cooperation between symbiotic rhizobia to overcome plant sanctions in the case of non-mutualist (e.g., which do not fix atmospheric nitrogen) or less-mutualist rhizobial partners (e.g., which have a limited nitrogenase functionality), but also may pave the way to 'free riders' or cheaters, which can be masked against sanctions directed toward the whole nodule, as shown on Bradyrhizobia (Denison, 2000).

The partnership between the rhizobium *Sinorhizobium meliloti* and legumes of genus *Medicago* (alfalfa and relatives) is one of the most investigated model systems of symbiotic nitrogen fixation. In this interaction, rhizobia are terminally differentiated into bacteroids inside root nodules (indeterminate nodules) and lose the ability to reproduce (Sprent, 2001; Gibson et al., 2008). However, indeterminate nodules also contain significant numbers of undifferentiated rhizobia, which can reproduce within the nodule tissues and are released after nodule senescence (Denison and Kiers, 2004). Therefore, a mixed rhizobial population inside the same nodule can be present, with both terminally differentiated and undifferentiated cells. The mixed (differentiated and undifferentiated) rhizobial population within the nodule may pose a question about the possibility of effective sanctions (Oono et al., 2009). Indeed, split-root experiments performed on *Sinorhizobium–Medicago* nodules did not support the hypothesis of sanctions over single nodules but indicate that a plant choice, occurring at some stage during infection, could be present (Gubry-Rangin et al., 2010). In this perspective, if single nodules are not sanctioned, we can hypothesize that the presence of multiple rhizobial strains in the same nodule with different level of mutualism (i.e. with different degrees of nitrogen fixation), would be more favored in indeterminate nodules (as those of *S. meliloti – Medicago*) than in determinate ones. Consequently, we can expect free-riding or cheating to be present in *S. meliloti–Medicago* symbiosis.

This work therefore aims at in providing experimental evidence on the presence of mixed single nodules and on the possibility of cheating/free riding to occur by non-mutualist strains in the *S. meliloti–Medicago* symbiosis. To fulfill this aim we evaluated:

(i) the presence of mixed nodules, which contain both the mutualist and the non-mutualist strain;

(ii) the differences in strains competitiveness in terms of nitrogen fixers nodules formed;

(iii) the fitness of the different strains in single and mixed nodules, in terms of number of bacterial cells present in nodules.

Two natural strains (plus the mutualist laboratory strain *S. meliloti* 1021) were selected for this study, a mutualist (BL225C) and a non-mutualist strain (AK83; Biondi et al., 2009) In particular, *S. meliloti* AK83 strain is able to form symbiotic nodules on alfalfa (*M. sativa*) but nitrogen fixation does not take place, either for the lack of a large gene set on the symbiotic megaplasmid pSymA (Galardini et al., 2011) and/or for the presence of an accessory plasmid which contains functions related to nitrogen fixation blocking (Crook et al., 2012; Price et al., 2015) Based on these premises, *S. meliloti* AK83 represents a good test strain for a free-riding/cheating behavior inside mixed root nodules.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, and Growth Conditions

The strains and plasmids used in this work are listed in **Table 1**. *S. meliloti* strains were cultured on solid or liquid tryptone yeast (TY) medium with 0.2 g/l $CaCO_3$, while *Escherichia coli* strains were grown in Luria Bertani (LB) medium, supplemented with antibiotics when necessary.

Preparation of Fluorescently Tagged Strains and CLSM Imaging

Strains were tagged with green fluorescent protein (GFP) or red fluorescent protein (RFP). Plasmids pHC60 (harboring a constitutively expressed GFP; Cheng and Walker, 1998) and pBHR mRFP (harboring a constitutively expressed RFP; Smit et al., 2005) were used for transformation of *E. coli* S17-1 cells. Transformants were selected by resistance to Tet (10 μg/ml). Positive clones were used for biparental conjugation

to a rifampicin-resistant derivative of *S. meliloti* AK83 and to *S. meliloti* 1021 (resistant to streptomycin). Conjugal transfer was performed as previously described (Pini et al., 2013b). Confocal imaging was performed on fresh cut nodules 4 weeks after inoculums, using an upright Leica Laser Scanning Confocal Microscope SP5 (Leica Microsystems, Germany) equipped with a 63 × oil immersion objective. Imaging of GFP and RFP were performed using 488-nm excitation of an argon laser line for GFP and 543-nm He/Ne for RFP.

Nodulation and Nitrogen Fixation Assays

Seedlings of *Medicago sativa* (cv. Pomposa) were sterilized in $HgCl_2$, repeatedly washed, and germinated in sterile plastic Petri dishes for 72 h in the dark and 48 h in the light at room temperature. For *in vitro* assays, seedlings were transferred in Petri dishes containing Buffered Nod Medium (Biondi et al., 2009) and 16 g/l of type A agar (Sigma–Aldrich). Plantlets were grown for an additional 3–5 days before inoculation with *S. meliloti* strains. For nodulation assays, strains were grown in liquid TY medium at 30°C for 48 h, then washed three times in 0.9% NaCl solution and resuspended to an $OD_{600\ nm}$ of 1.0. 1×10^7 cells (in single or a 1:1 ratio for competition experiments) were used (corresponding to 4×10^4 cells/cm^2). Cells were directly spread over the seedling root. Plates were pierced to let the plant grow outside, and transferred in a near-vertical position to a growth chamber maintained at 26°C with a 16-h photoperiod (100 microeinstein/m^2 s).

Microcosm-scale nodulation assays were performed in plastic pots containing approximately 400 g of native (unsterilized) sandy-clay soil from the university campus garden. Each pot was sown with four surface-sterilized seeds and after germination rhizobial strains (grown and washed as described for the *in vitro* assays) were inoculated into the pot to a final concentration of ca. 2×10^4 cells/g of soil, similar to the rhizobial estimates in agricultural soil. Plants were then maintained at 26°C with a 16-h photoperiod (100 microeinstein/m^2 s).

Nitrogen fixation rates were measured by the acetylene-reduction assay and expressed in nanomoles of produced ethylene per hour, per plant, as previously described in Muresu et al. (2008). All measures were taken 4 weeks after rhizobial inoculums, as commonly reported for *S. meliloti–Medicago* symbiotic tests (see for instance Yendrek et al., 2010; Pini et al., 2013a,b). Statistical analysis of data was performed with nonparametric Kruskal–Wallis test by Analyse-it software ver. 2.0 (Analyse-it Software, Ltd). Additional tests and figures were generated through Python scripts using the Numpy, Scipy (Van der Walt et al., 2011), Pandas (McKinney, 2010), Matplotlib (Hunter, 2007), and Seaborn (Waskom et al., 2015) libraries (**Supplementary Material S1**).

Estimation of Bacterial Loads in Nodules

Bacterial loads have been estimated by both a cultivation and a molecular (DNA-based) method. For cultivation, single (at least 10 days old, size >1 mm) nodules were excised from plants, surface sterilized with 0.1% NaHClO for 30″, washed three times in sterile distilled water, crushed and re-suspended in 100 µl of 0.9% NaCl sterile solution. Aliquots of serial dilutions and the third wash water (as control) were then plated on TY plates and incubated at 30°C for 48 h. The numbers of CFU were used for titration of viable and culturable cells. For the molecular method, bacterial DNA was extracted from surface-sterilized single nodules (as mentioned above) by using a fast protocol for plant DNA extraction (Edwards et al., 1991). Real-time PCR was performed by applying a previously published protocol (Trabelsi et al., 2009), based on the *S. meliloti* primer pair for core genome rpoE1 gene and on strain-specific primer pairs identified after comparative genome analysis (Galardini et al., 2011). In particular, for 1021, a strain-specific primer on Smc01419 gene (fw-5′-CGAGGAAGAGGTCCTGGAAT-3′, rv-5′-GACGCAGTCCTGCAACAGAT-3′) was designed, for AK83 strain on gene SINME_RS34005 (old locus tag Sinme_6912) (fw-5′-GATTTTCCGCGACTCTGAAG-3′, rv-5′-AGTCCGGTGTCAGATTCAGG-3′), and for BL225C on gene SinmeB 5863 (fw-5′GAAGCAGATGCTATCGGCAC-3′, rv-5′-TAAAACAGCACCACAGGCGAC-3′). All mentioned genes are single copy genes. Real-time (qPCR) was performed in an QuantStudio™ 7 apparatus (Applied Biosystems) programmed with the following temperature profile: 2 min 94°C, followed by 40 cycles composed by 15 s 94°C, 15 s 63°C, 30 s 72°C. Fluorescence data acquisition was done during the extension step at 72°C. A final melting curve was performed to check for product specificity. Reactions were performed in 10 µl final volume containing 5 µl of SYBR Green mix (Thermo Fisher Scientific Maxima SYBR Green/ROX qPCR Master Mix (2×), 0.5 µl ROX solution (included in the kit) and 10 pmol of primers. All reactions were done in triplicate. Standard curves with serial dilutions (1 ng–0.1 pg) of purified DNA from 1021, AK83, and BL225C were included. Data were analyzed with the QuantStudio™ 7 Flex System software (Applied Biosystems) using fast 96-Well Block (0.1 ml). PCR efficiency was calculated as in Bar et al. (2003). Bacterial cell number was estimated as genome copies (considering one genome copy per cell) present in a given DNA amount (genome sizes for strains are those derived from genome sequencing data (PRJNA42477 for BL225C, PRJNA41993 for AK83, PRJNA57603 for 1021).

Competition index (CI) was evaluated as log of the ratio of CFU or qPCR estimates between the two strains. by using a conventional metrics (Ellermeier et al., 2006; Crook et al., 2012) In particular, CI = log (titre of strain A recovered/titre of strain B recovered)/(titre of strain A inoculated/titre of strain B inoculated).

RESULTS

S. meliloti – M. sativa Symbiosis Allows a Large Number of Mixed Nodules

The common model of symbiotic interaction wrongly dictates the monoclonality of symbiotic rhizobia in root nodules, i.e. one nodule is colonized by one rhizobium strain only (Denison, 2000). Here, we provide evidence that both *in vitro* and in soil conditions, root nodules, formed after inoculating with a mixture of two rhizobial strains, contain both strains and that such nodules (when they include both mutualist and non-mutualist

strains) fix nitrogen at rates similar to root nodules formed by the sole mutualist strain.

First, we provided a quantitative estimation of the frequency of nodules with mixed infection in *M. sativa* plants under *in vitro* conditions. Single nodules were crushed, plated on TY plates and subjected to DNA extraction. Colonies grown on plates and extracted DNA from nodules were then analyzed by PCR with strain-specific primers (see MATERIALS AND METHODS). Results showed the presence of mixed nodules in all assays (20 plants with 3 nodules/plant for each competition). Titres estimated by plating and qPCR ranged from 10^1 to 10^6 cells/nodule (**Figures 1A,B**). In soil nodulation assays, performed

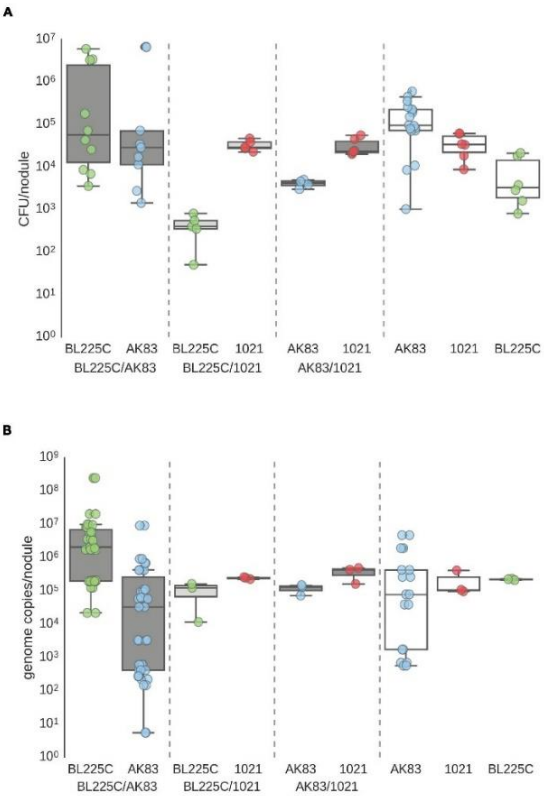


FIGURE 1 | Titres of *Sinorhizobium meliloti* on nodules (A) Viable titres estimates, (B) qPCR estimates. Viable cells/nodules in mixed nodules from competition experiments and single strain inocula (controls). Box-and-Whisker plots are from 3 to 20 single nodules, each collected from a different plant. Data are from plate experiment.

TABLE 1 | Strains and plasmids used.

Strains (or plasmids)	Species	Description	References
1021	*Sinorhizobium melloti*	Str^r derivative from strain 2011. Mutualist	Meade et al., 1982
AK83	*S. meliloti*	Lacks part of the microaerophilic gene set on pSymA-homolog megaplasmid. Non-mutualist	Galardini et al., 2011
BL225C	*S. meliloti*	Mutualist	Galardini et al., 2011
BM102	*S. meliloti*	Rif^r derivative of *S. meliloti* AK83. Non-mutualist	This work
BM267	*S. meliloti*	Rif^r derivative of *S. meliloti* BL225C. Mutualist	This work
BM685	*S. meliloti*	*S. meliloti* AK83 pBHR mRFP (Rif^r Tet^r). Non-mutualist	This work
BM687	*S. meliloti*	*S. meliloti* 1021 pBHR mRFP (Str^r Tet^r). Mutualist	This work
BM325	*S. meliloti*	*S. meliloti* AK83 pHC60 (Rif^r Tet^r). Non-mutualist	This work
BM257	*S. meliloti*	*S. meliloti* 1021 pHC60 (Str^r Tet^r). Mutualist	This work
BM286	*S. meliloti*	*S. meliloti* BL225C pHC60 (Rif^r Tet^r). Mutualist	This work
S17-1 λ*pir*	*Escherichia coli*	Tp^r, Sm^r, *recA*, thi, *hsdR*$^-$ M$^+$, RP4::2-Tc::Mu::Km::Tn7, λ*pir* lysogen	Simon et al., 1983
BM679	*E. coli*	*E. coli* S17-1 λ*pir* pBHR mRFP (Tet^r)	This work
BM266	*E. coli*	*E. coli* S17-1 λ*pir* pHC60 (Tet^r)	This work
pBHR mRFP		Constitutive expression of RFP, Tet^r	Smit et al., 2005
pHC60		Constitutive expression of GFP, Tet^r	Cheng and Walker, 1998

by inoculating garden soil with the rhizobial culture, confirmed the *in vitro* results, suggesting that in nature *S. meliloti* may indeed form mixed nodules with *M. sativa* (**Supplemental Figure S1**).

Molecular and plating data about the presence of mixed nodules were confirmed also by microscopy analysis on single *M. sativa* nodules colonized by GFP- and RFP-tagged derivatives of the strains (**Table 1**). Nodules with both strains present (AK83 and BL225C, as well as AK83 and 1021 or BL225C and 1021) were detected (**Figure 2**). Interestingly, both rhizobial strains were in non-differentiated form (normal bacterial-shape) in the meristematic part of the nodule and as intracellular bacteroid-shaped cells which can fix nitrogen (morphologically, thickened and long, Y-shaped cells surrounded

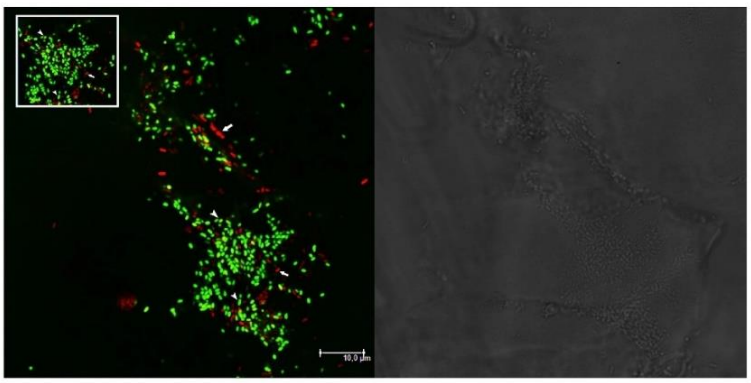


FIGURE 2 | Nodules may be colonized by different strains. Confocal laser scanning microscopy (CLSM) image of a portion of root nodule containing both BM687 (*S. meliloti* 1021 pBHR mRFP, in red) and BM286 (*S. meliloti* BL225C pHC60, in green) strains. Data from plate experiment. The bright field layer and the CLSM layer are shown. Arrows in the CLSM panel indicate both strains (in red and in green) colonizing the nodule. The inset highlights the central part of the CLSM image with BM687 and BM286 strains.

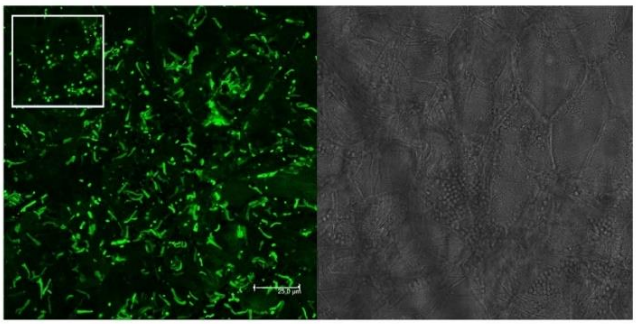


FIGURE 3 | Non-mutualist strain AK83 colonizes root nodules and undergo bacteroid differentiation. CLSM image of BM325 strain (*S. meliloti* AK83 pHC60) inside root nodules of *M. sativa* plants co-infected *in vitro* with strains BM325 and BM286. Y-shaped bacteroid-like cells of BM325 strain are abundant. The bright field layer and the CLSM layer are shown. The inset highlights a portion of the image were non-differentiated cells are present (as round-shaped cells).

by a membrane), also for the non-mutualist AK83 strain (**Figure 3**).

Finally, when the physiological efficiency of nodules was considered in terms of fixed nitrogen rates (acetylene-reduction assay), it emerged that the nitrogen fixation rate is not significantly reduced by the presence of the non-mutualist AK83 along with a mutualist strain (BL225C or 1021) (or, if it has a very faint effect, it is not statistically significant after Kruskal–Wallis test; **Figure 4**) in agreement with nodulation score and nodulation index which indicate that the presence of mixed infection in nodules, seems not to have a strong influence on plant growth (**Supplemental Figure S2**). From these data, we can support the hypothesis that the non-mutualist phenotype in terms of nitrogen fixation can be masked in mixed nodules by the

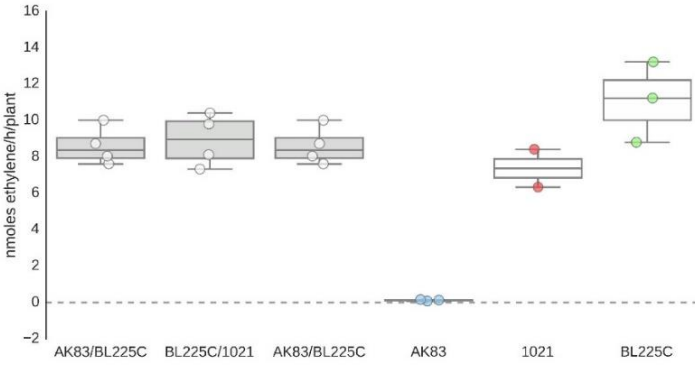


FIGURE 4 | Nitrogen fixation efficiency of nodules. Reduction assay results expressed in terms of nmoles ethylene/h/plant. Box-and-Whisker plots are from 3 to 4 nodules each from different plant.

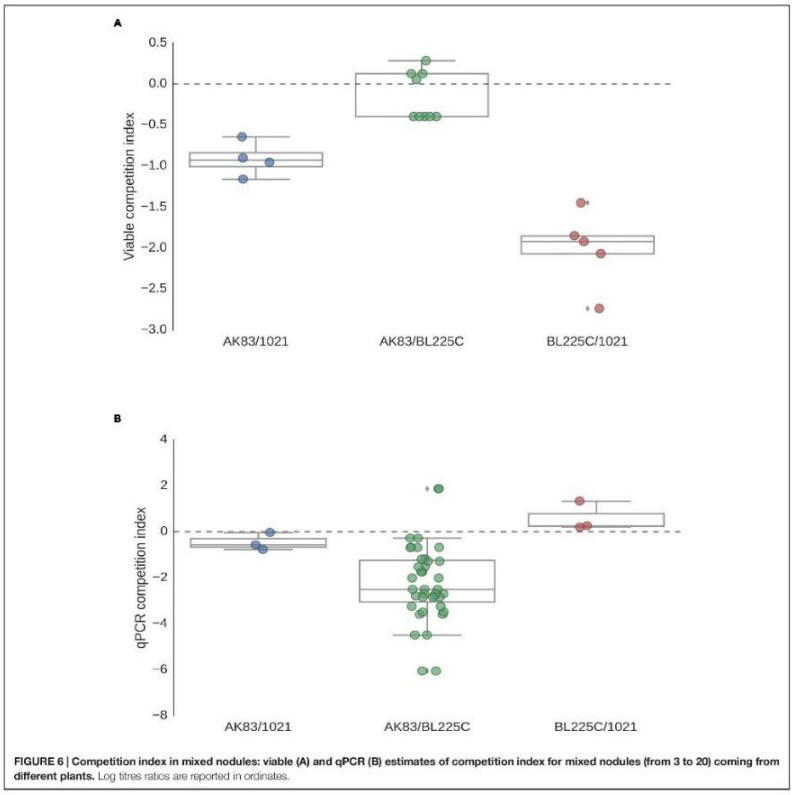


FIGURE 6 | Competition index in mixed nodules: viable (A) and qPCR (B) estimates of competition index for mixed nodules (from 3 to 20) coming from different plants. Log titres ratios are reported in ordinates.

This pattern of competition was also confirmed in an *in vivo* test on garden soil, with the mutualist strain (1021 and BL225C) being dominant in nodules with respect to the non-mutualist one (AK83) (Supplemental Table S1).

DISCUSSION

In this work we have, for the first time, generated evidences of widespread presence of mixed infection in nodules which contain cheater strains in the *Sinorhizobium–Medicago* symbiosis.

From the analysis of the colonization patterns, we found that nodules may host more than one strain. In several cases the non-mutualist strain was present with low number of cells inside the mixed nodules, but nodules in which it was present at high titres (similar or even higher than those of the mutualist strain) were also found. Plants nodulated by both mutualist and non-mutualist did not show an appreciable decrease in plant growth, nor root nodules showed a significantly large decrease in nitrogen fixation rates compared to plants nodulated by mutualist strains. These results support the hypothesis that, in nodules with mixed infection, the nitrogen

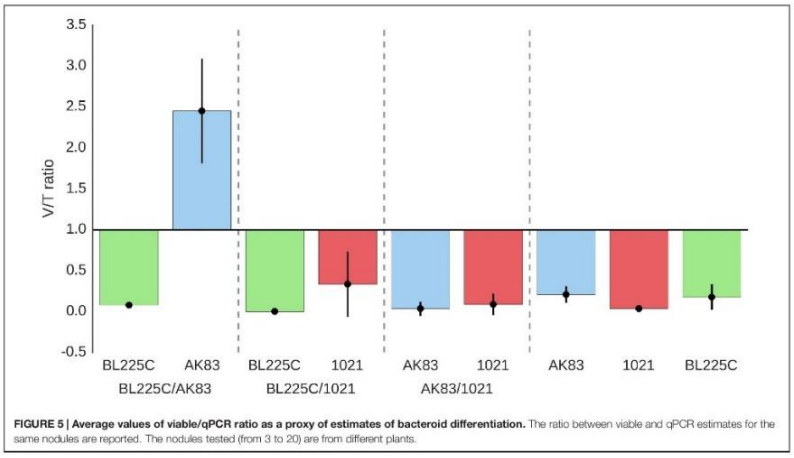


FIGURE 5 | Average values of viable/qPCR ratio as a proxy of estimates of bacteroid differentiation. The ratio between viable and qPCR estimates for the same nodules are reported. The nodules tested (from 3 to 20) are from different plants.

nitrogen fixation rate of the mutualist partner, thus avoiding plant sanctions toward the formation of mixed nodules. In particular, we can hypothesize that the good nitrogen fixation rates of mixed nodules can be due to an over-dominance of bacteroids of the mutualist strain with respect to the non-mutualist. Consequently, the number of rhizobial cells contained in single nodules was estimated.

The Non-mutualist Strain Eeficiently Colonizes Root Nodules

The amount of cells present in each nodule was considered as a proxy for fitness, as widely reported (see for instance (Kiers et al., 2003; Ratcliff et al., 2011). Both viable cells (after plating on TY plates) and total *S. meliloti* cells (estimated by qPCR) were counted on each nodule from single and mixed inocula. In most of the experiments, qPCR titres were higher (1–2 log) than viable titres (ratio < 1; **Figures 1A,B**). Although this difference might be due to technical reasons (difference in sensitivity between plating and qPCR), we cannot exclude that the consistently higher values of qPCR estimates may be due to the presence of the terminally differentiated bacteroids (which cannot be cultivated and contain multiple genome copies; Mergaert et al., 2006). Under this hypothesis, the ratio between viable and qPCR titres estimates was used to provide a possible proxy of the amount of strain differentiation within nodules (**Figure 5**). In the competition experiments between the mutualist BL225C and the non-mutualist AK83, viable/qPCR ratio was >1 for AK83, while approximated <1 for BL225C. These data let to hypothesize that the plant may favor the differentiation of the mutualist strain (BL225C). Such difference between the two competitors was not observed in the competition experiment between the two mutualist strains (as BL225C vs. 1021) as well as between 1021 and AK83 (values were below 1 for all strains). A similar higher colonization level of the mutualist strains (BL225C and 1021) with respect to strain AK83 was also observed in the soil experiment (**Supplemental Figure S1**). These results could be in agreement with a model of control over bacteroid differentiation operated by the plant (e.g., through NCR peptides), which favors the differentiation of mutualist strains, then maximizing the fixed nitrogen during symbiosis (Van de Velde et al., 2010; Price et al., 2015). However, it is still obscure whether some genetic determinant in the mutualist may positively affect the overall differentiation in the competition, similarly to the host range restriction peptidase (*hrrP*) present in the non-mutualist AK83 (Crook et al., 2012; Price et al., 2015).

However, we cannot exclude that estimates of bacteroid differentiation obtained with other methods, e.g., based on flow cytometry (Ratcliff et al., 2011) and evaluation of rhizobial cells in nodules at different stages could clarify the biological interpretation of the observed variability in qPCR estimates.

Competition indices (CI; **Figure 6**) showed that the non-mutualist strain AK83 was present, in many cases, at lower titres than the mutualist ones (BL225C and 1021, ratio <1). Then, in general, a lower fitness in competition of the non-mutualist with respect to the mutualist strain was detected, particularly in qPCR estimates (**Figure 6B**). However, nodules in which the non-mutualist colonized at titres higher than the mutualist strain were present, especially in the viable estimates of AK83 vs. BL225C competition (**Figure 6A**).

fixation phenotype of the mutualist strains (and consequently the effect on plant growth and, potentially, on plant fitness) is prevalent at the whole plant level. Consequently, the non-mutualist phenotype of some strains (as AK83) can be masked (against plant sanctions) by the fixed nitrogen provided by the mutualist strains. Then the non-mutualist strains can be protected against sanctions to ineffective (non-nitrogen fixing) nodules by the activity of the mutualist ones. After this consideration, we can support the hypothesis that a fraction of the rhizobial nodulating population could indeed be a cheater of the mutualist one (Denison, 2000; Denison and Kiers, 2004) or anyway may thrive, as free-rider, at the expenses of the mutualist strains. However, we cannot exclude that despite both viable and qPCR titres of nodule colonization indicate a dominance of the mutualist strains, a plant differential control over bacteroid formation could be present. In fact, different strains of *S. meliloti* (including AK83) seem to vary in their response to NCR peptides (Crook et al., 2012; Price et al., 2015), and consequently in the number of cells which do not terminally differentiate. In our experiments AK83 is forming bacteroid-shaped cells in *M. sativa* plants, even in mixed infection. This evidence indicates that the plant variety we used is producing NCR peptides able to differentiate AK83, then possibly limiting its infection as cheater. We can hypothesize that in other host plant species or other *M. sativa* varieties a different behavior of AK83 could be present, since AK83 strain harbors a *hrrP* homolog gene on pSINME01 plasmid (Price et al., 2015), which can selectively degrade some NCR peptides.

In general, the picture coming from the presented data is that the non-mutualist strains can have some space to thrive and multiply within the nodules of the root apparatus, without apparently reducing plant growth. The ability to multiply in the nodules (even if with reduced titres with respect to the mutualist strains) may allow to support the presence of a non-mutualist fraction in the rhizobia soil population (Nowak and Sigmund, 2002; Kiers et al., 2003). However, the consequences of this supposed cheating on plant fitness cannot be fully answered. In fact, although a "fitness" alignment has been reported between the rhizobia and the host plant (Friesen, 2012) and nodulation parameters have been shown to be correlated with plant biomass (Friesen, 2012), no measures of direct fitness (number of seeds, number of rhizobial cells released) were reported. Then, relevance of the presence of nodules with mixed infection containing non-mutualist strains on plant fitness has to be clarified. In fact, similar plant growth or nitrogen fixation rates could not imply directly similar number of seeds/offspring. Moreover, we cannot exclude that during nodule development the ratio of mutualist vs. non-mutualist strains may vary in relation to the differentiated and meristematic zone areas.

The impact of cheating on rhizobial evolution is hard to evaluate. Since rhizobial symbiosis is widespread in different bacterial taxa, even poorly related at the taxonomic level (cfr. the existence of alpha and beta-rhizobia), it is clear that symbiosis is conferring a fitness advantage for plant-associated bacteria (Remigi et al., 2014, 2016). However, the presence of other non-symbiotic ecological niches of rhizobia, as soil and plant endosphere, as well as the presence of genetic determinants modulating the competitiveness in different host plant varieties (Crook et al., 2012; Price et al., 2015) is puzzling in regard to the ecological and evolutionary relevance of cheating, with respect to the overall fitness of cheating strains. A strain behaving as cheater with one host plant genotype (variety, species, etc.) may be a good mutualist with a different plant genotype or having high fitness in a different environmental context (e.g., soil or rhizosphere). Moreover, the allowed presence of a supposed cheater might be based on other hitherto undetected phenotypes, different from nitrogen fixation, that could be beneficial for the plant. Indeed, it should be also kept in mind that, besides rhizobial cheaters, a high variety of endophytes of different taxonomy are regularly detected as co-occupants of nodules in many spontaneous legumes (Muresu et al., 2008; Zgadzaj et al., 2015). These endophytes may enter the nodule via Nod signaling from the true symbiotic strains (Zgadzaj et al., 2015) and may allow a fraction of non-symbiotic rhizobia to colonize root nodules formed by mutualistic rhizobia. However, the general role of root nodule non-symbiotic endophytes in affecting (positively or negatively) symbiotic nodule trophism is still unclear. Experiments in field conditions, evaluating both the nodulating and the non-nodulating fractions of the rhizobial population in competition experiments may allow to solve some of these questions.

AUTHOR CONTRIBUTIONS

AC performed competition experiments, *in vitro* symbiosis test produced recombinant strains and drafted the manuscript. EA, LM, SM performed CLSM experiments. AL and SM did phenotype testing on nodulated plants. AS and MZ performed nitrogen-fixation assays on nodules. MM settled the experimental protocol for competition experiments. MG performed the pangenome analysis of strains and prepared the figures. MB contributed in conceiving the idea. AM contributed in conceiving the idea and drafted the manuscript. All authors have read and approved the manuscript.

FUNDING

This work was supported by intramural funding of the University of Florence to AM and MB.

ACKNOWLEDGMENTS

The authors acknowledge the undergraduate students C. Gatsch, J. Lippi, R. Tomberli and S. Terzo for technical assistance to the *in vitro* symbiosis tests and for strain fluorescent labeling.

They are grateful to Federico Fontana for technical assistance in A.R.A. gas chromatography. We are grateful to Tom Bisseling for providing pBHR mRFP plasmid and to Yu-Xiang Jing for pHC60 plasmid and to Stefano Turillazzi for critical reading of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00835

FIGURE S1 | Number of copies of the genome/nodule of the qPCR estimates in the in vivo experiment. The plants were inoculated with the wild type strains. Measurements were performed on three nodules for each mixed inoculum.

FIGURE S2 | Effect of strain competition on host plants. (A) Nodulation index. Percentage of nodulated plants for single and mixed strains combinations. Values are means (±standard deviation) of three independent experiments, each involving at least 20 plants. **(B)** Nodulation score. Mean number of nodules/plant. Values indicate means ± standard deviation of number of three independent experiments, each involving at least 20 plants. The only significant pairwise contrast is that AK83 vs. BL225C-AK83 ($P<0.05$).

SUPPLEMENTAL MATERIAL S1 | Input data and code used to generate the figures.

REFERENCES

Bar, T., Ståhlberg, A., Muszta, A., and Kubista, M. (2003). Kinetic outlier detection (KOD) in real-time PCR. *Nucleic Acids Res.* 31:e105. doi: 10.1093/nar/gng106

Biondi, E. G., Tatti, E., Comparini, D., Giuntini, E., Mocali, S., Giovannetti, L., et al. (2009). Metabolic capacity of *Sinorhizobium (Ensifer) meliloti* strains as determined by phenotype microarray analysis. *Appl. Environ. Microbiol.* 75, 5396–5404. doi: 10.1128/AEM.00196-09

Carelli, M., Gnocchi, S., Fancelli, S., Mengoni, A., Paffetti, D., Scotti, C., et al. (2000). Genetic diversity and dynamics of *Sinorhizobium meliloti* populations nodulating different alfalfa varieties in Italian soils. *Appl. Environ. Microbiol.* 66, 4785–4789. doi: 10.1128/AEM.66.11.4785-4789.2000

Cheng, H. P., and Walker, G. C. (1998). Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by *Rhizobium meliloti*. *J. Bacteriol.* 180, 5183–5191.

Crook, M. B., Lindsay, D. P., Biggs, M. B., Bentley, J. S., Price, J. C., Clement, S. C., et al. (2012). Rhizobial plasmids that cause impaired symbiotic nitrogen fixation and enhanced host invasion. *Mol. Plant Microbe Interact.* 25, 1026–1033. doi: 10.1094/mpmi-02-12-0052-r.

Denison, R. F. (2000). Legume sanctions and the evolution of symbiotic cooperation by rhizobia. *Am. Nat.* 156, 567–576. doi: 10.1016/j.cub.2011.06.018

Denison, R. F., and Kiers, E. T. (2004). Lifestyle alternatives for rhizobia: mutualism, parasitism, and forgoing symbiosis. *FEMS Microbiol. Lett.* 237, 187–193. doi: 10.1016/j.femsle.2004.07.013

Dwivedi, S. L., Sahrawat, K. L., Upadhyaya, H. D., Mengoni, A., Galardini, M., Bazzicalupo, M., et al. (2015). Advances in host plant and rhizobium genomics to enhance symbiotic nitrogen fixation in grain legumes. *Adv. Agron.* 129, 1–116

Edwards, K., Johnstone, C., and Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* 19:1349. doi: 10.1093/nar/19.6.1349

Ellermeier, C. D., Hobbs, E. C., Gonzalez-Pastor, J. E., and Losick, R. (2006). A three-protein signaling pathway governing immunity to a bacterial cannibalism toxin. *Cell* 124, 549–559. doi: 10.1016/j.cell.2005.11.041

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

Galardini, M., Mengoni, A., Brilli, M., Pini, F., Fioravanti, A., Lucas, S., et al. (2011). Exploring the symbiotic pangenome of the nitrogen-fixing bacterium *Sinorhizobium meliloti*. *BMC Genomics* 12:235. doi: 10.1186/1471-2164-12-235

Gibson, K. E., Kobayashi, H., and Walker, G. C. (2008). Molecular determinants of a symbiotic chronic infection. *Annu. Rev. Genet.* 42, 413–441. doi: 10.1146/annurev.genet.42.110807.091427

Gubry-Rangin, C., Garcia, M., and Béna, G. (2010). Partner choice in *Medicago truncatula*–*Sinorhizobium* symbiosis. *Proc. Biol. Sci.* 277, 1947–1951.

Hunter, J. D. (2007). Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* 9, 90–95. doi: 10.1109/MCSE.2007.55

Kiers, E. T., Rousseau, R. A., and 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., Rousseau, R. A., West, S. A., and Denison, R. F. (2003). Host sanctions and the legume-rhizobium mutualism. *Nature* 425, 78–81. doi: 10.1038/nature01931

Mårtensson, A., Brutti, L., and Ljunggren, H. (1989). Competition between strains of *Bradyrhizobium japonicum* for nodulation of soybeans at different nitrogen fertilizer levels. *Plant Soil* 117, 219–225. doi: 10.1007/BF02220715

McKinney, W. (2010). "Data structures for statistical computing in python," in *Proceedings of the 9th Python in Science Conference*, Austin, TX, 51–56.

Meade, H. M., Long, S. R., Ruvkun, G. B., Brown, S. E., and Ausubel, F. M. (1982). Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn5 mutagenesis. *J. Bacteriol.* 149, 114–122.

Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., et al. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the Rhizobium-legume symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* 103, 5230–5235. doi: 10.1073/pnas.0600912103

Moawad, M., and Schmidt, E. (1987). Occurrence and nature of mixed infections in nodules of field-grown soybeans (*Glycine max*). *Biol. Fertil. Soils* 5, 112–114. doi: 10.1007/BF00257644

Muresu, R., Polone, E., Sulas, L., Baldan, B., Tondello, A., Delogu, G., et al. (2008). Coexistence of predominantly nonculturable rhizobia with diverse, endophytic bacterial taxa within nodules of wild legumes. *FEMS Microbiol. Ecol.* 63, 383–400. doi: 10.1111/j.1574-6941.2007.00424.x

Nowak, M. A., and Sigmund, K. (2002). Biodiversity: bacterial game dynamics. *Nature* 418, 138–139. doi: 10.1038/418138a

Oono, R., Denison, R. F., and Kiers, E. T. (2009). Controlling the reproductive fate of rhizobia: how universal are legume sanctions? *New Phytol.* 183, 967–979. doi: 10.1111/j.1469-8137.2009.02941.x

Pini, F., Frage, B., Ferri, L., De Nisco, N. J., Mohapatra, S. S., Taddei, L., et al. (2013a). The DivJ, CbrA and PleC system controls DivK phosphorylation and symbiosis in *Sinorhizobium meliloti*. *Mol. Microbiol.* 90, 54–71. doi: 10.1111/mmi.12347

Pini, F., Spini, G., Galardini, M., Bazzicalupo, M., Benedetti, A., Chiancianesi, M., et al. (2013b). Molecular phylogeny of the nickel-resistance gene nreB and functional role in the nickel sensitive symbiotic nitrogen fixing bacterium *Sinorhizobium meliloti*. *Plant Soil* 1–13. doi: 10.1007/s11104-013-1979–1973

Price, P. A., Tanner, H. R., Dillon, B. A., Shabab, M., Walker, G. C., and Griffitts, J. S. (2015). Rhizobial peptidase HrrP cleaves host-encoded signaling peptides and mediates symbiotic compatibility. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15244–15249. doi: 10.1073/pnas.1417797112

Ratcliff, W. C., Underbakke, K., and Denison, R. F. (2011). Measuring the fitness of symbiotic rhizobia. *Symbiosis* 55, 85–90. doi: 10.1098/rspb.2009.2072

Remigi, P., Capela, D., Clerissi, C., Tasse, L. N., Torchet, R., Bouchez, O., et al. (2014). Transient hypermutagenesis accelerates the evolution of legume endosymbionts following horizontal gene transfer. *PLoS Biol.* 12:e1001942. doi: 10.1371/journal.pbio.1001942

Remigi, P., Remigi, P., Zhu, J., Young, J. P., and Masson-Boivin, C. (2016). Symbiosis within symbiosis: evolving nitrogen-fixing legume symbionts. *Trends Microbiol.* 24, 63–75. doi: 10.1016/j.tim.2015.10.007

Sessitsch, A., Hjemba, P. K., Hardarson, G., Akkermans, A. D. L., and Wilson, K. J. (1997). Measurement of the competitiveness index of *Rhizobium tropici* strain CIAT899 derivatives marked with the gusA gene. *Soil Biol. Biochem.* 29, 1099–1110. doi: 10.1016/S0038-0717(97)00002-3

Simon, R., Priefer, U., and Puhler, A. (1983). A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Nat. Biotech.* 1, 784–791. doi: 10.1038/nbt 1183-784

Smit, P., Raedts, J., Portyanko, V., Debellé, F., Gough, C., Bisseling, T., et al. (2005). NSP1 of the GRAS protein family is essential for rhizobial Nod factor-induced transcription. *Science* 308, 1789–1791. doi: 10.1126/science. 1111025

Sprent, J. I. (2001). *Nodulation in Legumes*. London: Royal Botanic Gardens, Kew.

Trabelsi, D., Pini, F., Aouani, M. E., Bazzicalupo, M., and Mengoni, A. (2009). Development of real-time PCR assay for detection and quantification of *Sinorhizobium meliloti* in soil and plant tissue. *Lett. Appl. Microbiol.* 48, 355–361. doi: 10.1111/j.1472-765X.2008.02532.x

Van der Walt, S., Chris Colbert, S., and Varoquaux, G. (2011). The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng.* 13, 22–30.

Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., et al. (2010). Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* 327, 1122–1126. doi: 10.1126/science.1184057

Waskom, M., Botvinnik, O., Hobson, P., Warmenhoven, J., Cole, J. B., Halchenko, Y., et al. (2015). *seaborn: v0.6.0 (June 2015)*. doi: 10.5281/zenodo.19108

Yendrek, C. R., Lee, Y. C., Morris, V., Liang, Y., Pislariu, C. I., Burkart, G., et al. (2010). A putative transporter is essential for integrating nutrient and hormone signaling with lateral root growth and nodule development in *Medicago truncatula*. *Plant J.* 62, 100–112. doi: 10.1111/j.1365-313X.2010.04134.x

Zgadzaj, R., James, E. K., Kelly, S., Kawaharada, Y., de Jonge, N., Jensen, D. B., et al. (2015). A legume genetic framework controls infection of nodules by symbiotic and endophytic bacteria. *PLoS Genet.* 11:e1005280. doi: 10.1371/journal.pgen.1005280

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

Chapter 3

3.1 Role and regulation of ACC deaminase gene in *Sinorhizobium meliloti*: is it a symbiotic, rhizospheric or endophytic gene?

In this work, we focused on the gene *acdS*, located in the dispensable genome of a large variety of *S. meliloti* strains. According to previous studies, the gene, which is supposed to be involved in the degradation of the ethylene precursor (ACC), seems to be linked also to the improvement of symbiotic performance and competitiveness of rhizobial strains. Indeed, the gene is located in *S. meliloti* on the symbiotic megaplasmid (pSymA), supporting a hypothesis of relationships with the symbiosis. However, no data are present on its actual functional role in relation with the symbiotic performance, as well as to the modulation of ethylene production by the plant.

frontiers in Genetics

ORIGINAL RESEARCH
published: 30 January 2017
doi: 10.3389/fgene.2017.00006

Role and Regulation of ACC Deaminase Gene in *Sinorhizobium meliloti*: Is It a Symbiotic, Rhizospheric or Endophytic Gene?

Alice Checcucci[1†], Elisa Azzarello[2], Marco Bazzicalupo[1], Anna De Carlo[3], Giovanni Emiliani[3], Stefano Mancuso[2], Giulia Spini[2†], Carlo Viti[2] and Alessio Mengoni[1*]

[1] Department of Biology, University of Florence, Sesto Fiorentino, Italy, [2] Department of Agri-food Production and Environmental Science, University of Florence, Florence, Italy, [3] Consiglio Nazionale delle Ricerche (CNR), Istituto per la Valorizzazione del Legno e delle Specie Arboree, Florence, Italy

OPEN ACCESS

Edited by:
Naoki Osada,
Hokkaido University, Japan

Reviewed by:
Birgit Mitter,
Austrian Institute of Technology, Austria
Akinori Yabuki,
Japan Agency for Marine-Earth Science and Technology, Japan

***Correspondence:**
Alessio Mengoni
alessio.mengoni@unifi.it

†Present Address:
Giulia Spini,
Institut of Microbiology, Catholic University of the Sacred Heart, Piacenza, Italy

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: 25 November 2016
Accepted: 13 January 2017
Published: 30 January 2017

Citation:
Checcucci A, Azzarello E, Bazzicalupo M, De Carlo A, Emiliani G, Mancuso S, Spini G, Viti C and Mengoni A (2017) Role and Regulation of ACC Deaminase Gene in Sinorhizobium meliloti: Is It a Symbiotic, Rhizospheric or Endophytic Gene?. Front. Genet. 8:6. doi: 10.3389/fgene.2017.00006

Plant-associated bacteria exhibit a number of different strategies and specific genes allow bacteria to communicate and metabolically interact with plant tissues. Among the genes found in the genomes of plant-associated bacteria, the gene encoding the enzyme 1-aminocyclopropane-1-carboxylate (ACC) deaminase (*acdS*) is one of the most diffused. This gene is supposed to be involved in the cleaving of plant-produced ACC, the precursor of the plant stress-hormone ethylene toning down the plant response to infection. However, few reports are present on the actual role in rhizobia, one of the most investigated groups of plant-associated bacteria. In particular, still unclear is the origin and the role of *acdS* in symbiotic competitiveness and on the selective benefit it may confer to plant symbiotic rhizobia. Here we present a phylogenetic and functional analysis of *acdS* orthologs in the rhizobium model-species *Sinorhizobium meliloti*. Results showed that *acdS* orthologs present in *S. meliloti* pangenome have polyphyletic origin and likely spread through horizontal gene transfer, mediated by mobile genetic elements. When *acdS* ortholog from AK83 strain was cloned and assayed in *S. meliloti* 1021 (lacking *acdS*), no modulation of plant ethylene levels was detected, as well as no increase in fitness for nodule occupancy was found in the *acdS*-derivative strain compared to the parental one. Surprisingly, AcdS was shown to confer the ability to utilize formamide and some dipeptides as sole nitrogen source. Finally, *acdS* was shown to be negatively regulated by a putative leucine-responsive regulator (LrpL) located upstream to *acdS* sequence (*acdR*). *acdS* expression was induced by root exudates of both legumes and non-leguminous plants. We conclude that *acdS* in *S. meliloti* is not directly related to symbiotic interaction, but it could likely be involved in the rhizospheric colonization or in the endophytic behavior.

Keywords: *Sinorhizobium meliloti*, ACC deaminase, ethylene, *acdS*, nitrogen sources, endophytic colonization, rhizosphere

INTRODUCTION

Plant-bacteria interactions have been studied since long time as reciprocal beneficial association (symbiosis), as neutral interaction (commensalism), and as pathogenic interaction. Despite many details are known on the molecular bases of all the above-mentioned interactions (Lugtenberg et al., 2002) a number of genes present in the genome of plant-associated bacteria is still under debate and a complete explanation of the various mechanisms used by plant-associated bacteria is still lacking. Recent analyses suggested the presence of a core set of genes in plant-associated bacterial genomes (Pini et al., 2011), which include genes related to transport, regulation, sugar metabolism, etc. However, many plant-associated bacteria exhibit several additional genes, related to the different type of interaction they have (e.g., nitrogenase for symbiotic rhizobia). One of the mostly diffused additional genes among rhizospheric and endophytic bacteria is that encoding the enzyme 1-aminocyclopropane-1-carboxylate (ACC) deaminase, referred to as *acdS* (Nascimento et al., 2014; Singh et al., 2015). Biochemically, ACC deaminase is able to degrade the precursor of ethylene biosynthesis, ACC, into ammonium and α-ketobutyrate (Honma and Shimomura, 1978). Ethylene is working as a plant hormone and affects all stages of plant development and growth (Deikman, 1997), mainly in relation with biotic and abiotic stresses (Abeles and Heggestad, 1973). The ACC deaminase structural gene (*acdS*) has been found in many rhizosphere bacteria, in symbiotic rhizobia, in bacterial endophytes, in fungi and in the genomes of several plants, as *Arabidopsis* (Singh et al., 2015).

The presence of ACC deaminase in plant-associated bacteria, has been interpreted as a way to use the additional nitrogen source (represented by ACC), consequently decreasing the amount of ACC available by the plant for the production of the phytohormone ethylene (Glick et al., 1998; Holguin and Glick, 2003; Prigent-Combaret et al., 2008; Gamalero and Glick, 2015). The reduction of ethylene production by the plant may have positive effect over colonization of plant tissue by bacteria. Indeed, ethylene is known to have an inhibitory effect on rhizobial infection, limiting formation and number of nodule for plants and the root growth (Nukui et al., 2006). The decrease of ethylene emission may increase root system development (Penrose and Glick, 2003) and enhance nutrients and water uptake (Reid and Renquist, 1997), thus allowing a higher number of symbiotic nodules to be formed on host plant root system. Moreover, endophytic plant growth promoting bacteria (as *Burkholderia phytophirmans* PsJN, *Pseudomonas fluorescens* YsS6, and *Pseudomonas migulae* 8R6) are less effective when their *acdS* orthologs are deleted (Sun et al., 2009; Ali et al., 2012; Nascimento et al., 2014). In the nitrogen-fixing symbiont *Mesorhizobium loti acdS* gene has been shown to be transcribed inside root nodules, suggesting an involvement in the symbiotic process.

Phylogenetic analyses suggested that horizontal gene transfer (HGT) has played a strong role in *acdS* spreading within taxonomic groups (Blaha et al., 2006; Nascimento et al., 2012; Lemaire et al., 2015). On the other hand, recently (Nascimento et al., 2014), a detailed phylogenetic reconstruction has been performed, showing that *acdS* orthologs are preferentially vertically inherited along the bacterial phylogeny. However, due to the scattered occurrence in the same species, HGT events can still be supposed, at least at the species or genus level, and selective advantages conferred to strains has to be clarified. In particular, comparative genomic analyses have shown that *acdS* orthologs are part of the dispensable genome fraction in species as the model symbiotic rhizobium *Sinorhizobium meliloti*. In *S. meliloti*, a previous genome analysis suggested *acdS* as one of the genes which may explain different symbiotic phenotypes among strains (Galardini et al., 2011). However, no experimental indication of its role in the symbiotic performance was reported.

Previous works demonstrate the presence and the correlation of a regulatory region upstream to *acdS* gene belonging to the *lrp* family (leucine-responsive regulatory gene like), called *acdR* (Grichko and Glick, 2000; Ma et al., 2003b) in *Pseudomonas putida* UW4, *Rhizobium leguminosarum* bv.viciae 128C53K (Grichko and Glick, 2000), *Azospirillum lipoferum* 4B and most other *acdS*+ *Proteobacteria* (Prigent-Combaret et al., 2008), confirming that usually in *Proteobacteria* the regulatory genes are close to the genes they regulate. However, its mode of regulation (in relation with *acdS* and its promoter) is not totally clarified, especially in relation with the symbiotic partner plants.

In this work, we aimed at define the evolution and the functional profile of *acdS* and its regulatory gene *acdR* in the model plant symbiont *S. meliloti*. Our study showed that HGT has played a strong role in shaping *acdS* phylogeny in *S. meliloti*, suggested additional roles, not related with ethylene modulation and symbiosis, which may have selected its presence in the dispensable genome fraction of *S. meliloti* (Nascimento et al., 2014) and allowed to confirm common trends on the evolution of modules of regulatory interactions in bacteria (Babu et al., 2006).

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

The strains and plasmids used in this work are listed in **Table 1** and Supplementary Table S1. In particular a collection of *S. meliloti* strains from different geographical areas was used (Carelli et al., 2000; Roumiantseva et al., 2002, 2014; Talebi Bedaf et al., 2008; Trabelsi et al., 2009, 2010). *S. meliloti* strains were cultured on solid or liquid TY medium with 0.2 g/liter $CaCO_3$, or Vincent Minimal Medium (VMM, or Rhizobium Defined Medium, RDM), while *Escherichia coli* strains were grown in Luria Bertani (LB) medium, supplemented with antibiotics when necessary.

Detection of *acdS* Gene, Genomic Context, Analysis, and Phylogenetic Reconstruction

The presence of *acdS* orthologs in a collection of 133 *S. meliloti* strains was performed by PCR amplification on crude lysates using the two sets of primers and the PCR conditions described in Duan et al. (2009). *S. meliloti* 1021 was used as negative control, while *S. meliloti* AK83 was used as positive controls. Agarose gel electrophoresis on 1.5% TAE buffer and ethidium bromide staining (10 mg/l) was used for visualization of amplification products on an UV transilluminator. Positive amplification

TABLE 1 | Strains and plasmids used in the work.

Strains (or plasmids)	Description	References
S. meliloti 1021, BM 678	Str^r derivative from strain 2011	Meade et al., 1982
S. meliloti AK83, BM 493	Lacks part of the microaerophilic gene set on pSymA-homolog megaplasmid	Galardini et al., 2011
E. coli S17-1 λpir	Tp^r, Sm^r, *recA*, tni, *hsdR*$^-$M$^+$, RP4::2-Tc::Mu::Km::Tn7, λpir lysogen	Simon et al., 1983
BM 193	*E. coli* S17-1 λpir + expression vector lac promoter regulation pSRK-Km (Km^r)	L. Ferri, unpublished
BM 637	*E. coli* S17-1 λpir + pSRK-Km (Km^r) with *acdS* gene from AK83 strain	This work
BM 261	*S. meliloti* 1021 pSRK- Km (Km^r)	Pini et al., 2014
BM 641	*S. meliloti* 1021 + pSRK- *acdS* AK83 (Km^r)	This work
BM 634	*E. coli* S17-1 λpir +promoterless vector GFP pOT2 (Gm^r)	This work
BM 668	*S. meliloti* 1021 pOT2 (Gm^r)	This work
BM 689	*E. coli* S17-1 λpir pOT2 + *acdS* promoter (144 bp) (Gm^r)	This work
BM 690-144 bp	*S. meliloti* 1021 pOT2 + *acdS* promoter (144 bp) (Gm^r)	This work
BM 674	*E. coli* S17-1 λpir pOT2 + *acdS* promoter + regulator (620 bp) (Gm^r)	This work
BM 697-620 bp	*S. meliloti* 1021 pOT2 + *acdS* promoter+ regulator (620 bp) (Gm^r)	This work

BM; Laboratory ID code, stands for "Bazzicalupo Marco"

products from two strains, representative of the collection (BO21CC, 2B13) were cloned into pGEM®-T Easy Vector Systems (Promega) following manufacturer's instructions and sequenced for confirmation of *acdS* presence.

Orthologs of *acdS* and *acdR* were retrieved from GenBank database running a blast search over Rhizobiaceae (taxid:82115) non-redundant nucleotide database on 2016-05-16 by using *acdS* gene from AK83 (Sinme_5642) and *acdR* from AK83 (Sinme_5643) as query sequence. The alignment of aminoacidic sequences was performed with ClustalW (Goujon et al., 2010). For phylogenetic reconstruction, a Model Test was run on aligned sequences to choose the best substitution model (Supplementary Data 1). The model with the lowest Bayesian Information Criterion (Schwarz, 1978; Nei and Kumar, 2000) was chosen for running Maximum Likelihood phylogenetic reconstruction (Anisimova and Gascuel, 2006). Robustness of dendrograms topology was inferred by running 500 bootstrap replicates. All steps of the phylogenetic reconstruction were performed with MEGA 7 software (Kumar et al., 2016). Alignments and nexus files of the phylogenetic reconstructions are included as Supplementary Data 1.

Preparation of Cloning Vectors and Transformation of Strains

The *acdS* gene from AK83 (Sinme_5642) strain was cloned into pSRK-Km vector under lac-promoter (Khan et al., 2008) and firstly used for transformation of *E. coli* S17-1 cells. Transformant cells were selected for resistance to Km (10 μg/ml), and positive clones were used for biparental conjugation to *S. meliloti* 1021 (resistant to streptomycin, 200 μg/ml). Conjugal transfer was performed as previously described (Pini et al., 2014). Gene expression was induced by treating *in vitro* plantlets inoculated with IPTG (at concentration of 0.23 mM). The promoter (144 bp fragment) and the promoter in association with transcriptional regulator (620 bp fragment) from BL225C strain was cloned into the promoter-less vector pOT2 containing GFP-uv (green fluorescent protein) as reporter gene (Karunakaran et al., 2005). Recombinant vectors were used for transformation of *E. coli* S17-1 cells selected for resistance to Tc (10 μg/ml), then the positive clones were used for biparental conjugation to *S. meliloti* 1021.

ACC Deaminase Assay

Permeabilized cells were obtained from 5 ml overnight liquid cultures after harvesting cells by centrifugation, washing the pellet with 0.9% NaCl solution. Cell permeabilization was performed by adding by 600 μl of 100 mM Tris HCl pH 8.5 and 30 μl of toluene and vortexing for 30 s. After 1 h incubation at 4°C, lysed cells were centrifuged at 12000 rpm for 10 min and toluene was removed. The permeabilized cell suspensions were used for total protein content determination with Bradford reagent (Sigma-Aldrich) and enzymatic assays. ACC deaminase activity was quantified on crude cell extracts by measuring the amount of α-ketobutyrate produced by the deamination of ACC, as previously described by Honma and Shimomura (1978) and Penrose and Glick (2003).

In vitro Symbiosis Assays

Seedlings of *Medicago sativa* (cv. Pomposa) were sterilized in $HgCl_2$, repeatedly washed, and germinated in sterile plastic Petri dishes for 72 h in the dark and 48 h in the light at room temperature. For *in vitro* assays, seedlings were transferred in Petri dishes containing Buffered Nod Medium (Ehrhardt et al., 1992) and 16 g/l of type A agar (Sigma-Aldrich). Plantlets were grown for an additional 3 to 5 days before inoculation with *S. meliloti* 1021, *acdS*-derivative and the parental strains. For nodulation assays, strains were grown in liquid TY medium at 30°C for 48 h with antibiotics if necessary, then washed three times in 0.9% NaCl solution and resuspended to an OD_{600} nm of 1.0. Then, aliquots of 1×10^7 cells were used, as previously described (Pini et al., 2013, 2014). Cells were directly spread over the seedling root. Plates were kept in a growth chamber maintained at 26°C with a 16-h photoperiod (100 microeinstein m^{-2} s^{-1}) for 40 days.

Ethylene Measurement

Ten *M. sativa* (cv. Pomposa) germinated seeds (treated as previously described) were singularly sown in 120 ml glass vials containing 30 ml of Buffered Nod Medium (Biondi et al., 2009) and 16 g/l of type A agar (Sigma-Aldrich). Seedlings were grown for additional 2 days before inoculation with *S. meliloti* strains (1021 with the empty pSRK vector, 1021 *acdS*-derivative and the parental strain). The strains were grown in liquid TY medium at 30°C for 48 h, washed in 0.9% NaCl solution and resuspended to an OD 600 nm of 0.5 as previously described. Each vial was then inoculated with 500 μl of bacterial suspension (corresponding to $\sim 5 \times 10^4$ cells).

The vials, sealed with Teflon septum and crimped with aluminum caps, were kept in a growth chamber at 23°C ± 1, under a 16 h photoperiod and 60 μmol m^{-2} s^{-1} photosynthetically active radiation provided by cool-white fluorescent lamps. The ethylene accumulation was detected at 30 or 60 days post inoculation (dpi).

Ethylene concentrations in the headspace were determined using an ultra-sensitive ETD-300 photo-acoustic laser spectrophotometer (Sensor Sense B.V., Nijmegen, The Netherlands. http://www.sensor-sense.nl) in combination with a gas handling system. In brief, the detector consists of a CO_2 laser and a photo-acoustic cell, in which the gas is detected. The detector is able to detect on-line about 300 parts per trillion by volume of ethylene within a 5-s time scale. The gas handling was performed by a valve control box (type VC-6, Sensor Sense B.V., Nijmegen, the Netherlands), designed for measuring up to six sampling cuvettes per experiment. In this experiment, the valve control box allowed automated sampling of ethylene accumulated into vials at a flow rate of 3 l h^{-1} and its transport to the ETD-300 alternately, in succession for 15 min for each cuvette.

The air control was sourced completely from the compressed air source and was measured to contain less than 0.001 μl l^{-1} ethylene. Statistical analysis of data has been performed with one-way ANOVA and Tukey *post-hoc* comparison by using Past software (Hammer et al., 2001).

Nodule Colonization Measurement

Estimation of bacterial loads in nodules in single and mixed inocula has been performed with a Real Time PCR method (Checcucci et al., 2016). In brief, single nodules of the same size (~1 mm in length) were excised from plants, surface sterilized with 0.1% NaHClO and crushed for the DNA extraction. Real Time PCR was performed with the *acdS* specific primers and *S. meliloti* specific primers (Trabelsi et al., 2009) *acdS* specific primers (fw-5′- TGAATTGTGTCGTCATCCAG -3′, rv-5′- CTGTCGGCGCCCATCAGTTT-3′) were designed with Primer3 software (http://primer3.sourceforge.net/)on the basis of *acdS* gene from AK83 strain (Sinme_5642) from position 371379 nt to position 371479 nt

Phenotype Microarray Experiments

Phenotype microarray (PM) experiments were performed to investigate the metabolic functions carried out by AcdS. *S. meliloti* 1021 pSRK- Km (BM261) and *S. meliloti* 1021 + pSRK-*acdS* AK83 (BM641) strains were assayed by PM technology (Biolog) using microplates PM3, PM6, PM7, and PM8, which test different nitrogen and peptides compounds sources. PM data were analyzed to compare the activities of 1021 wild type strain with those of its derivative expressing *acdS*. Strains were grown at 30°C on TY agar plates for 2 days and then, colonies were picked with a sterile cotton swab from the agar surface and suspended in 15 ml of NaCl 0.8% until a cell density of 81% transmittance (OD_{600} = 0.1) was reached on a Biolog turbidimeter. The inoculation fluids for PM panels was prepared adding 2 ml of each cell suspension and 240 μl of dye MixA 100x (Biolog) to 22 ml of M9 media depleted from carbon source and enriched with IPTG 1 mM (necessary for the activation of the promoter of pSRK vector and the expression of the gene). Then the inoculation fluid was transferred to the microplates (100 μl per well). All PM microplates were incubated at 28°C in an OmniLog reader, and changes of color in the wells were monitored automatically every 15 min. Readings were recorded for 96 h, and data were analyzed using OminoLog PM software which generated a time course curve for tetrazolium color development.

Positive PM results were then confirmed by inspecting the growth of *S. meliloti* 1021 pSRK- Km (BM261) and *S. meliloti* 1021 + pSRK- acdS AK83 (BM641) strains on VVM medium containing formamide as sole nitrogen source after 24 h at 30°C.

Formamidase Activity Assay

Formamidase activity present in crude permeabilized cells was performed by using the Berthelot reaction with a colorimetric determination of ammonium (Anderson and Ingram, 1993) as described in Skouloubris et al. (1997) with minor modifications. Aliquots of 50 μl of cell extracts in 100 mM Tris-HCl pH 8.5 buffer were mixed with 100 μl of 100 mM formamide solution in the same buffer. After 30 min incubation at room temperature 400 μl of salycilate-citrate-nitroprusside solution was added and incubated for 15 min followed by the addition of 400 μl of the alkaline hypochlorite reagent. After 1 h incubation sample absorbance at 655 nm was read. Blank samples were prepared by boiling cell extracts 20 min prior to the addition of formamide. Ammonia released was determined from a standard curve. Formamidase activity was expressed in units corresponding to the degradation of 1 μmol of formamide min^{-1} mg^{-1} protein.

Promoter Activation and Regulation Patterns

To investigate *acdS* gene promoter activation patterns, putative *acdS* promoter region (144 bp) and the upstream region including also its putative transcriptional regulator (620 bp) were cloned in the promoter-less vector pOT2 (Karunakaran et al., 2005). Firstly, the plasmids were used for transformation of *E. coli* S17-1 cells, then transformant cells were selected for resistance to Gentamicin (10 μg/ml), and positive clones were used for biparental conjugation to *S. meliloti* 1021. Conjugal transfer was performed as previously described (Pini et al., 2014). Recombinant *S. meliloti* 1021 strains [*S. meliloti* 1021 pOT2 + *acdS* promoter (144 bp) (BM 690), and *S. meliloti* 1021 pOT2 + *acdS* promoter+ regulator (620 bp) (BM 697)] were grown on TY plates and the induction tests were performed by

suspending a single colony in tubes with M9, M9 supplemented with NH_4Cl (10 g/l) and M9 supplemented with ACC (5 mM)., and in tubes with 300 µl 0.9% NaCl solution and 200 µl of root exudates.(Ogawa and Long, 1995). After incubation for 3 h, cultures were placed in a microtiter plate and GFP-uv gene expression was measured on a microplates reader (Tecan Infinite 200 PRO, Tecan, Switzerland).

Root Exudates Production

Seedlings of eight leguminous and not leguminous plant species (*M. sativa* cv. Pomposa, *Phaseolus vulgaris*, *Lens culinaris*, *Arabidopsis thaliana*, *Nicotiana tabacum*, *Daucus carota*, *Rafanus sativus*, and *Lepidium sativum*) were used for the production of root exudates. *M. sativa* seeds were sterilized in $HgCl_2$, repeatedly washed in sterilized water, and germinated in sterile plastic Petri dishes for 72 h in the dark and 48 h in the light at room temperature with 2–3 ml of sterile ddH_2O. *A. thaliana* seeds were sterilized for 1′ in EtOH 70% and in a solution of Bleach 10%, ddH_20 90%, and Triton X-100 0.1% for 10′. The seeds were then repeatedly washed and germinated in Magenta waving containing M&S based Medium. The other seeds were sterilized similarly to *A. thaliana*, but germinated in sterile plastic Petri dishes for 5–6 days in the light in a growth chamber maintained at 26°C with a 16-h photoperiod (100 microeinstein m^{-2} s^{-1}). All the plantlets were then grown for 6–7 days, and then transferred over a filter of Whatman paper, in 13 ml tubes with 10 ml of sterile ddH_2O. Each tube contained approximately the same amount of root biomass. After 2 weeks of incubation in the growth chamber, the exudates were extracted, filtered and stored at −80°C.

Confocal Imaging

Plantlets were germinated and grown on BNM medium plates as previously described (see Symbiotic and nodulation assays). One-week-old plantlets were placed on a sterilized microscope slide spread with a thin layer of BNM Agar Medium and inoculated with 50 µl (corresponding to 4×10^4 cells/cm^2 on the glass slide surface) from an overnight culture of 1021 pOT2 + *acdS* promoter (144 bp) (BM 690), and 1021 pOT2 + *acdS* promoter+ regulator (620 bp) (BM 697) strains grown in TY medium and washed three times with 0.9% NaCl solution. Images were taken using an upright Leica laser-scanning confocal microscope SP5 (Leica Microsystems Wetzlar GmbH, Germany)

RESULTS

Occurrence and Phylogeny of *acdS* in *S. meliloti*

We inspected the presence of *acdS* genes in a collection of 133 *S. meliloti* strains from different geographical areas (Supplementary Table S1) Thirty-one strains (22.6% of total) gave positive amplification. The percentages of positive strains varied from 0% (Tunisia) to 44% (Italy). Kazakhstan and Iran strains collections both showed 13% of positive strains. In the Kazakhstan collection, the six positive strains were distributed among different host plants, viz. *M. lupulina*, *M. falcata*, *M. trautvetteri*, *Melilotus* sp.

The phylogenetic analysis, based upon *acdS* orthologs found among Rhizobiaceae, highlighted the presence of two main clades of the *acdS* orthologs in *S. meliloti* (**Figure 1A**), distributed within *R. leguminosarum* strains. The genomic context analysis of the *acdS* orthologs from the two clades, results showed the presence of Mobile Genetic Element (MGE), as transposases and integrases in the close proximity of *acdS* gene, in every *S. meliloti* strain (**Figure 1B**).

Function and Control of Acds in *S. meliloti*

To shed light on the functional roles of *acdS* in *S. meliloti*, *acdS* gene from AK83 (Sinme_5642) was cloned under *lac* promoter and introduced into *S. meliloti* 1021 strain [producing 1021 + pSRK- *acdS* AK83 (BM641), see **Table 1**]. This strain showed a positive ACC deaminase activity under IPTG induction (Supplementary Table S2). Then, to investigate the functional role of *acdS* in *S. meliloti* and its putative involvement in the reduction of plant ethylene production, the ethylene accumulation of the host plants infected by the recombinant and wild type strains was measured (**Figure 2**). No statistically significant difference in ethylene produced by the host plants was detected and all the samples tested showed values similar to the control (not infected plant) ($0.5 < p$-values > 0.005).

We then tested the possible role of *acdS* expression in the symbiotic performances and competitiveness of *S. meliloti*. In the symbiotic test, both single and mixed inocula did not show significant differences in the percentage of nodulated plants (**Figure 3A**) as well as in the overall rhizobial colonization of root nodules (p-value < 0.5) (**Figure 3B**). Concerning the competitiveness inside nodule and the capability of colonization, tough both strains were detected, the empty vector 1021 strain [1021 pSRK- Km (BM261)] showed for most of the nodules a higher titre than its *acdS*-expressing derivative 1021 + pSRK-*acdS* AK83 (BM641) (**Figure 3C**) (mean 7.14×10^2, standard deviation 1.05×10^3, median 1.74×10^2). This result suggested that *acdS* expression does not allow to better compete in nodule colonization, but that in our experimental setup the expression of *acdS* under lac promoter possibly reduced the growth and or the differentiation abilities inside root nodule tissue.

To evaluate additional metabolic abilities conferred by *acdS*, Phenotype Microarrays were performed. We tested a total of 384 different nitrogen and peptide sources using plates PM3, PM6, PM7, and PM8, on BM261 and its *acdS*-expressing derivative 1021 + pSRK- *acdS* AK83. Results showed almost complete identical behavior for the two strains but, surprisingly, 1021 + pSRK- *acdS* AK83 strain induced with IPTG displayed higher metabolic activity in formamide than the control 1021 pSRK-Km (BM261). 1021 + pSRK- *acdS* AK83 strain was indeed able to grow better on formamide as sole nitrogen source, then its parental BM261 strain (in **Figure 4**, the results obtained in Biolog Plate PM3), as confirmed by the growth on VVM medium (Supplementary Table S3). However, 1021 + pSRK- *acdS* AK83 strain induced with IPTG did not show higher formamidase activity than BM261 strain (data not shown), hampering to evaluate the biochemical basis of the detected growth difference.

As shown in **Figure 1B**, in the genomic context analysis, a quite conserved region upstream to *acdS* gene in every *S. meliloti*

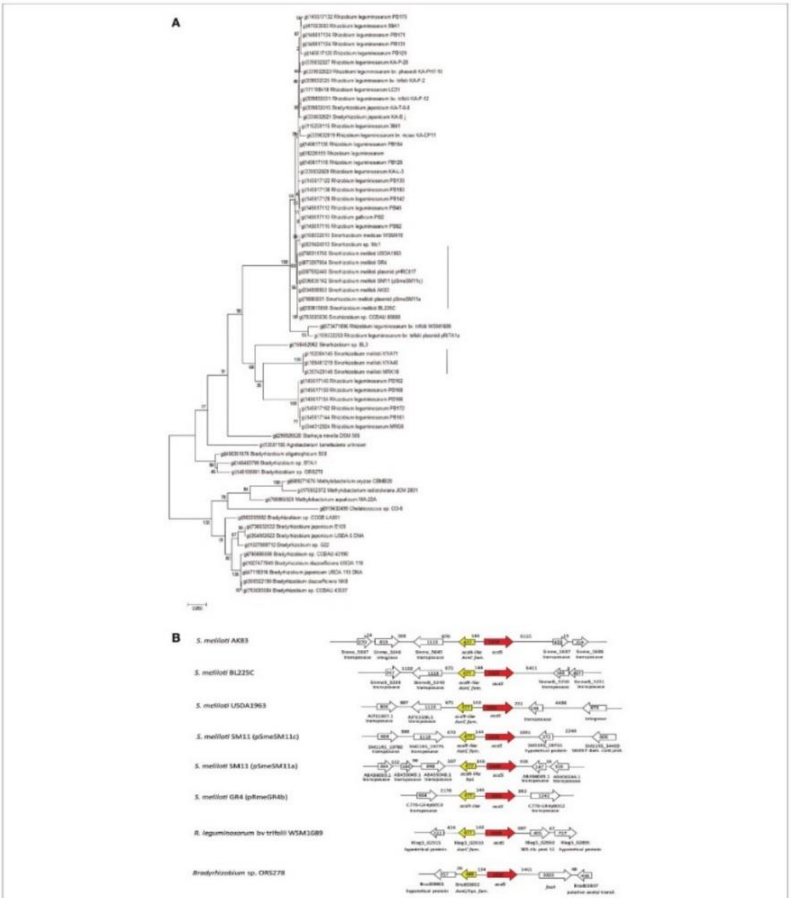


FIGURE 1 | Phylogeny of the ACC deaminase gene cassette. (A) Maximum Likelihood phylogenetic reconstruction of *acdS* gene sequences The phylogenetic analysis is based upon *acdS* orthologs found in Rhizobiaceae. The dendreogram highlight the presence of two clades of the gene in *S. meliloti*., vertical lines **(B)** Genomic context of the *acdS* orthologs from the two clades of *acdS* orthologs in *S. meliloti* strains and from *Bradirhizobium* and *Rhizobium leguminosarum* strains. *acdS* region map pointed out the presence of MGE (Mobile Genetic Elements) close to *acdS*. The length (bp) of genes and intergenic regions is indicated, as well as ORF orientation (using arrows). The GTR+G model has been chosen for the reconstructions after model test evaluation (Supplementary Data 1).

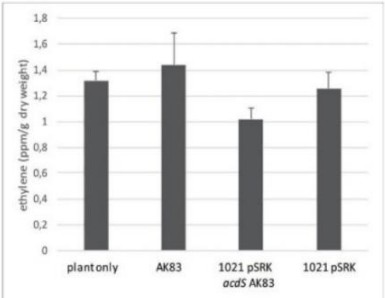


FIGURE 2 | Effect of rhizobial inoculants on plant ethylene production. Ethylene accumulation measurements in terms of ppm of plant dry weight are reported. The final measurements were performed 60 days after inoculum. Reported values indicate average from 5 replicates. Error bars indicate standard deviation. No significant differences (p-values < 0.5) between inoculants have been detected (one- way ANOVA).

strains is present. The region is composed by an intergenic spacer (the putative promoter, 144 bp long), downstream to an open reading frame, in opposite orientation with respect to *acdS*, annotated as a putative leucine-responsive regulator (lrpL/AsnC family) (476 long, here called *acdR* in agreement with previous naming; Grichko and Glick, 2000; Prigent-Combaret et al., 2008). We considered as putative promoter the DNA fragment between the translation start codons of *acdS* gene and that of the putative regulator (Ma et al., 2003a). Previous works have described that in other species, as *R. leguminosarum* bv. *viciae* 128C53K, the *acdR* gene (a putative leucine-responsive regulator, named also as *lrpL*) is required for the expression of *acdS* (Ma et al., 2003a, 2004) and that the only presence of both *acdS* and *lrpL* genes can allow the expression of ACC deaminase (Stiens et al., 2006). Since a putative LRP box is present in the 144 fragment (5′-AAGCAAAATTAGAGA-3′ at 62 nt from the AcdS start codon) we wanted to investigate *acdS* regulation in relation to the presence of the putative *acdR* gene. Consequently, we cloned the sole putative promoter (1021 pOT2 + *acdS* promoter (144 bp), BM690 in **Table 1** and the entire region, 620 bp long, [putative promoter and the *acdR* gene, 1021 pOT2 + *acdS* promoter+ regulator (620 bp), BM697 strain in **Table 1**] from strain AK83 into the promoter-less vector pOT2 and tested reported gene activation (GFP-uv) in *S. meliloti* 1021. The strain with the sole promoter (1021 pOT2 + *acdS* promoter (144 bp) (BM690) showed transcriptional activity higher (p-values < 0.0001) than the strain with the putative leucine-responsive regulator [1021 pOT2 + *acdS* promoter+ regulator (620 bp), BM697] in M9 medium supplemented with both ACC and ammonium as sole nitrogen source (**Figure 5**), suggesting a negative control of the transcriptional regulator AcdR toward *acdS*.

We then investigated which conditions may allow the release of the repression by AcdR. considering that M9 mineral medium supplemented with ACC as sole nitrogen source was tested but no activity was detected. (**Figure 5**). Then, to observe if plant proximity may be a factor allowing to induce gene expression, 1021 pOT2 + *acdS* promoter + regulator (620 bp) strain was spread close to the roots of *M. sativa* plantlets and visualized by fluorescence microscopy. Results showed that *M. sativa* roots are able to induce promoter activation (**Figure 6**). No fluorescence was observed with the 1021 pOT2 (BM668, empty vector) (data not shown). Finally, to further quantitatively evaluate the level of promoter activation and understand if the activation may be specific of the symbiotic host plant (*Medicago* spp.), we incubated BM697 and BM690 strains in presence of root exudates (Ogawa and Long, 1995). The results are showed in terms of ratio between the level of activation of 1021 pOT2 + *acdS* promoter + regulator (620 bp) (BM697) and of 1021 pOT2 + *acdS* promoter (144 bp) (BM690) (**Figure 7**). The results showed that most of all tested root exudates induced promoter activation (ratio > 0.4), in particular those of *M. sativa, L. culinaris, Rafanus sativum, P. vulgaris, and Lepidum sativum*, allowing BM697 to restore the level of GFPuv expression of the strain with the sole promoter 1021 pOT2 + *acdS* promoter (144 bp) (BM690). Interestingly, the data highlight root exudate from *Daucus carota* showed statistically significant differences with all the other tested root exudates (ratio > 1).

DISCUSSION

The gene encoding ACC deaminase (*acdS*) is considered to be important for bacteria plant interaction mainly since it is considered to lower the level of ethylene produced by the plant (Gamalero and Glick, 2015; Singh et al., 2015). However, in relation to rhizobial plant symbiosis only few data were present. In particular, for *Mesorhizobium loti-Lotus japonicum* association *acdS* activity has been shown to be present inside mature root nodules, in relation to NifA2 control (Nukui et al., 2006). For *S. meliloti*, past works have shown an increase in competitiveness of an engineered strain containing *acdS* genes with respect to its wild-type counterparts (Ma et al., 2004) and an increase in host plant growth (*M. lupulina*) when infected with an ACC deaminase-overproducing *S. meliloti* strain (Kong et al., 2015). However, no details on the frequency of *acdS* in *S. meliloti* strains, as well as on its regulation and functional role were present. Moreover, in these works *acdS* from other species (*P. putida* and *R. leguminosarum* bv. *viciae*) were used, consequently no indications of the actual role of *S. meliloti* native *acdS* were reported.

Distribution and Evolutionary Pattern of *acdS* Gene in *S. meliloti*

In this work, we have shown that *acdS* genes have undergone extensive horizontal transfer events in *S. meliloti*. In particular, the analysis of a collection of 133 strains coming from Iran, Kazakhstan, Tunisia, and Italy showed that *acdS* is present in ca. 22% only of strains, thus confirming earlier reports indicating that *acdS* is part of the dispensable genome fraction in *S. meliloti* (Galardini et al., 2011). Some differences in the

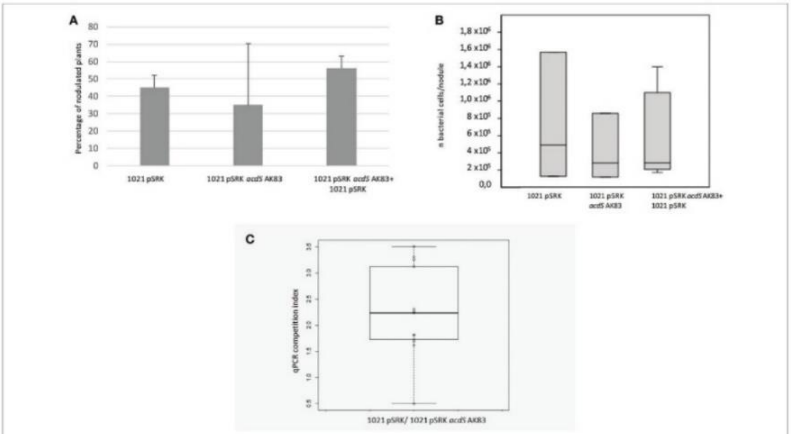


FIGURE 3 | *acdS* expression does not confer competitive advantage in symbiosis. (A) Nodulation efficiency and **(B)** competitiveness insides nodules of 1021 overexpressing *acdS* from AK83 strain compared to the parental one are reported. **(A)** The percentage of nodulated plants is evaluated on a set of 50 infected plants for each strain (20 for singular inoculum and 30 for the mixed inoculum). **(B)** Boxplots reports the rhizobial titres inside root nodules taken from 13 independent measures (different nodules from different plants) for each inoculant strain. **(C)** qPCR competition index (CI) to evaluate capability in the colonization of the nodules, reported as log of the ratio of qPCR estimates of rhizobial titres (number of rhizobial cell in the nodule, evaluated for each strain involved in the competition) between the two strains inoculated in the same individual plant (the parental strain 1021 with the empty vector, and strain 1021 overexpressing *acdS* from AK83 strain). The boxplot reports the data from 13 nodules from different co-inoculated plants.

in the percentage of *acdS* harboring strains from the different geographical areas were found. Even if a two-way PERMANOVA indicate a statistical significance of the geographical area (data not shown), on the basis of actual data related, we cannot indicate if such difference may be due to stochastic effects (linked to the composition of the collection) or to the host plants used for strain isolation. Indeed, Tunisian strains were isolated with *M. truncatula* only (Trabelsi et al., 2010), while Iran and Italian strains were isolated on *M. sativa* plants only, though from different cultivars (Carelli et al., 2000; Talebi Bedaf et al., 2008). The six Kazakhstan strains containing *acdS* were isolated from different hosts (either *Medicago* and *Melilotus*), but numbers are not adequate to provide a statistical evaluation of possible host plant preference. Sampling experiments performed with different host plant species in controlled conditions are needed to clarify if strains carrying *acdS* are advantaged during symbiosis with specific plant species.

The considerable level of horizontal spreading of *acdS* is present also in other rhizobia. A search over Integrated Microbial Genome Database (IMG, (Markowitz et al., 2013) showed that *acdS* is present in 94% of the completely sequenced *Bradyrhizobium* strains and in the 33% of *R. leguminosarum* strains, confirming that also in such genera/species is part of the dispensable genome fraction (data not shown). Indeed such horizontal spreading has been highlighted in the whole Bacteria domain (Hontzeas et al., 2005; Blaha et al., 2006; Nascimento et al., 2014). In Nascimento et al. (2014) *S. meliloti acdS* sequences appeared split into two clades, the one containing strains AK83, BL225C and SM11, the other strains KYA40 and KYA71. We confirmed here the occurrence of these two clades for *acdS* in *S. meliloti*, suggesting that within *S. meliloti* pangenome *acdS* may have originated from different transfer events. The detected presence of Mobile Genetic Elements (MGE) close to *acdS* in *S. meliloti* may suggest recent HGT events of *acdS* in *S. meliloti*. This genome organization in *S. melioti* appears to be quite similar to that of other bacterial species, as the well-investigated strain *P. putida* UW4 (Grichko and Glick, 2000; Li and Glick, 2001). Interestingly, MGE are not present at close distance in other *Rhizobiaceae* (as *B. japonicum* or *R. leguminosarum*), suggesting that the spreading of *acdS* in *S. meliloti* should have been more recent. This hypothesis is supported by the evidence that in the genomes of *S. meliloti* strains, *acdS* is present on the symbiotic megaplasmids (related to pSymA of strain 1021), which are known to be of relatively recent origin and have undergone large structural rearrangements, especially by MGE movements (Galardini et al., 2013). Moreover, it is quite relevant to notice

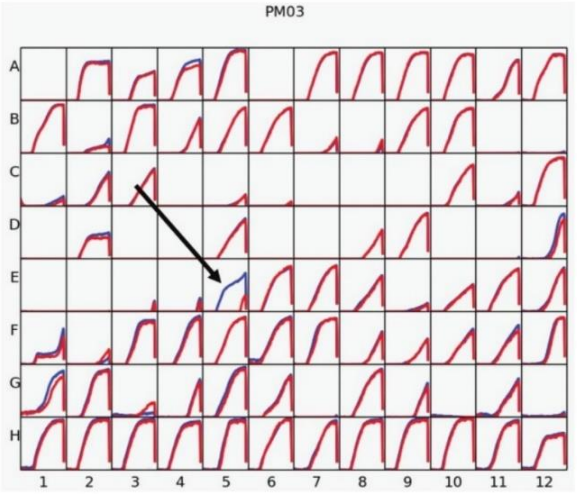


FIGURE 4 | Phenotype Microarray (PM). The results of PM analysis for nitrogen sources on Biolog Plate 3 in Rm1021 overexpressing *acdS* from AK83 strain compared to the parental one is reported. Measures were taken incubating strains with IPTG for inducing *acdS* gene expression. The red kinetic curves represent the strain 1021, while the blue one report the metabolic activity of 1021 pSRK *acdS* AK83. The black arrow underline the results obtained for the activity of the strains in presence of formamide.

that in *S. meliloti* SM11 strain, two *acdS* genes were found and they located into two different replicons, one in pSME11a and the other one in pSME11c (related to pSymA of the model strain 1021) (Schneiker-Bekel et al., 2011). Finally, we found upstream to several *acdS* orthologs the presence of the putative regulator *acdR*, highlighting a conservation of the gene cassette (Supplementary Figure S1).

Functions and Regulation of *acdS*

Previous studies indicated that expression of ACC deaminase increases nodulation ability of *S. meliloti* toward *M. sativa* (Ma et al., 2004). Our results did not provide clear evidences of an effect on increase in competitiveness of the *acdS* expressing strain with respect to the parental one, neither as percentage of nodulated plants, nor as overall nodule colonization. Moreover, the competition index based on qPCR estimation in *M. sativa* nodules, showed on the contrary that the expression of *acdS* under lac promoter reduced the colonization of the nodules to the advantage of the parental strain. Of course, we cannot *a priori* exclude that other plant varieties and testing conditions may allow to detect differences. It is in fact known that symbiotic test may provide variable results, depending on the strain used and on the plant genotypes (Crook et al., 2012). However, we can hypothesize that ACC deaminase expression did not provide a considerable advantage to the bacterium in the symbiotic interaction.

Concerning the potential ability of ACC deaminase in the reduction of plant ethylene production (Glick, 2005), our results did not support this conclusion. However, we cannot exclude that, because slightly less ethylene (differences were not statistically significant) was present in the plants inoculated with the *acdS* expressing strains, an effect could be detected by analyzing a higher biomass of plants or in different experimental conditions (e.g., with plants challenged with a stressing agent, as salt or heavy-metals). Indeed, in other systems (e.g., *M. loti*), ACC deaminase may lower plant ethylene levels, but only locally (Murset et al., 2012), suggesting then that on the overall plant (as in our conditions) effects could be minimized. Interestingly, Phenotype Microarray data showed a surprising phenotype of the *acdS* expressing strain, which was able to use formamide and Ile-Pro dipeptide as sole nitrogen source. This result led us to formulate a hypothesis of a role of ACC deaminase as scavenger of unusual nitrogen sources (in the rhizosphere and/or in the plant endosphere). Indeed, a

formamide concentration dependent growth was shown for the recombinant *acdS* expressing strain, though also the parental strain *S. meliloti* 1021 showed some ability to grow on formamide as sole nitrogen source. Such a metabolic hypothesis on ACC deaminase role in rhizobia in the scavenging of unusual nitrogen sources could allow to explain the presence of *acdS* in some non-mutualist rhizobial strains (Checcucci et al., 2016). In other words, ACC deaminase activity could allow some strains to better perform in rhizosphere and endosphere colonization because of increase nutrient availability, then also to explain the increased nodulation ability found in some rhizobial species (Ma et al., 2004). Indeed, *acdS* among rhizobia is not ubiquitous and different results in relation to the nodulation and symbiosis have been highlighted in different species, as *B. japonicum* (Murset et al., 2012). However, the presence of *acdS* in the dispensable genome fraction and its polyphyletic pattern of evolution in *S. meliloti*, strongly suggest that the conferred advantage is only strain-specific (depending on the genomic background of the single strain) or that a balancing selection is acting on the *S. meliloti* population, reducing in some way the possible fitness advantages of *acdS*-harboring strains. Furthermore, the presence of the gene in a large number of non-symbiotic rhizospheric and endophytic bacterial species (Gamalero and Glick, 2015) support the hypothesis of its role in the colonization of the rhizosphere and endosphere.

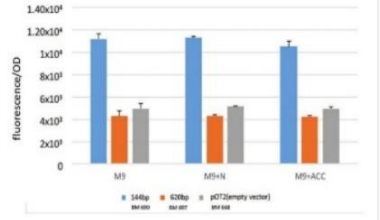


FIGURE 5 | Transcriptional control of *acdS* gene. The activation of gene expression of strains BM690 (144 bp), BM697 (620 bp), and BM688 (1021 pOT2-empty vector) in terms of GFP fluorescence/OD in M9 medium, M9 supplemented with ammonium and M9 supplemented with ACC is reported. Values indicate average from 3 replicates with standard deviation. Significant statistical differences with ANOVA TEST were found (p-values < 0.0001).

The hypothesis of a main involvement of *acdS* in the broadening of rhizobial metabolic abilities (e.g., for the colonization of rhizosphere) and not in symbiosis was finally supported by the results of *acdS* gene regulation. We showed that the *acdR*-like gene is potentially acting as repressor of *acdS* expression. This result is in contrast with what was found for the orthologs present in *P. putida* (previously *Enterobacter cloacae* UW4), where a positive regulation mediated by AcdR was present (Grichko and Glick, 2000; Nukui et al., 2006), as well as with the data reported for *M. loti* where NifA is involved in promoting *acdS* expression (Nukui et al., 2006). Interestingly, in *S. meliloti* the repression operated by AcdR was not released by the presence of ACC in the medium, but by the presence of root and in particular by the incubation with root exudates from either *M. sativa* and other species (both leguminous and not leguminous plants), again in support of additional roles (not just the ACC degradation) of *acdS* in *S. meliloti*. In particular, these data let to hypothesize that most of the root exudates are able to activate the promoter and the transcription. Moreover, *D. carota* results suggest the presence of other molecules in root exudates, which act as positive modulators.

Finally, the different regulatory pattern of *acdS* in *S. meliloti* with respect to *P. putida* UW4, even in presence

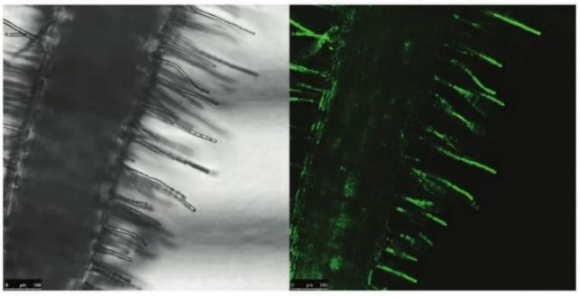


FIGURE 6 | Root proximity activates *acdS* gene expression. Fluorescence confocal image of a portion of *M. sativa* root infected by BM697 (620 bp) strain. GFP expression is in close proximity of hairy roots.

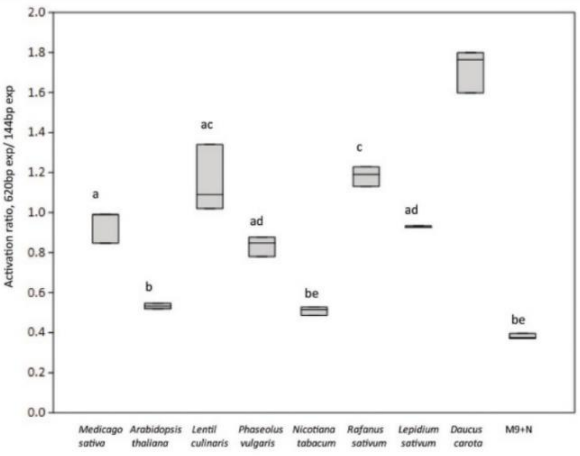


FIGURE 7 | Roots exudates from different plant species activates *acdS* gene expression. Boxplots report the activation ratio between the expression (exp) of the entire cassette with respect to the sole promoter (BM897/BM890 strains (620 bp exp/144 bp exp) is reported. Different letters indicate significantly different values (one-way ANOVA, $p < 0.05$).

of corresponding orthologs, is an interesting example of previous findings on the evolution of regulatory interactions in bacteria, where a higher conservation of effectors than of regulatory schemes is observed in different bacterial species (Babu et al., 2006; Galardini et al., 2015).

On the overall, the presented results strongly suggested that *acdS* spread in *S. meliloti* pangenome in relation to the colonization of plant roots more than to the symbiotic interaction. Consequently, we can hypothesize that *acdS* may be linked to an increase of fitness in non-symbiotic host plant species. The involvement of *acdS* in such non-symbiotic role may have contributed to the expansion of *S. meliloti* ecological niche. Indeed, pSymA megaplasmid is showing other non-symbiotic genes, as *nreB* (Pini et al., 2014, 2015), suggesting additional roles of pSymA, other than symbiosis and nitrogen-fixation. Indeed, comparative genomic analyses showed the pSymA megaplasmid to be a hot spot for structural variations (Galardini et al., 2015) Consequently, we can speculate that pSymA is undergoing evolutionary changes that partly can mirror those already occurred in the pSymB chromid (diCenzo et al., 2014), where additional genes integrated into an ancient dispensable plasmid, increasing the metabolic load and allowing to expand the ecological niche of *S. meliloti*, and ultimately leading of the formation of a non-dispensable element, the chromid.

AUTHOR CONTRIBUTIONS

AC designed the work, performed most the analyses, provided interpretation of data, contributed in conceiving the work and drafted the manuscript. EA, SM, AD, GE, GS, and CV contributed analyses and provided interpretation of data. AM and MB conceived the work, provided interpretation of data and drafted the work. All authors contributed critically revised manuscript and gave the final approval for publication.

FUNDING

This work was performed under partial funding assigned to AM (University of Florence, Finanziamento progetti strategici ricerca di base 2014).

ACKNOWLEDGMENTS

The authors are grateful to Elisa Gamalero for the critical reading of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fgene.2017.00006/full#supplementary-material

REFERENCES

Abeles, F. B., and Heggestad, H. E. (1973). Ethylene: an urban air pollutant. *J. Air Pollut. Control Assoc.* 23, 517–521. doi: 10.1080/00022470.1973.10469798

Ali, S., Charles, T. C., and Glick, B. R. (2012). Delay of flower senescence by bacterialendophytes expressing 1-aminocyclopropane-1-carboxylate deaminase. *J. Appl. Microbiol.* 113, 1139–1144. doi: 10.1111/j.1365-2672.2012.05409.x

Anderson, J. M., and Ingram, J. S. I. (eds.). (1993). "Colorimetric determination of ammonium," in *Tropical Soil Biology and Fertility: A Handbook of Methods, 2nd Edn* (Wallingford: CAB International), 73–74.

Anisimova, M., and Gascuel, O. (2006). Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst. Biol.* 55, 539–552. doi: 10.1080/10635150600755453

Babu, M. M., Teichmann, S. A., and Aravind, L. (2006). Evolutionary dynamics of prokaryotic transcriptional regulatory networks. *J. Mol. Biol.* 358, 614–633. doi: 10.1016/j.jmb.2006.02.019

Biondi, E. G., Tatti, E., Comparini, D., Giuntini, E., Mocali, S., Giovannetti, L., et al. (2009). Metabolic capacity of *Sinorhizobium* (Ensifer) *meliloti* strains as determined by phenotype microarray analysis. *Appl. Environ. Microbiol.* 75, 5396–5404. doi: 10.1128/AEM.00196-09

Blaha, D., Prigent-Combaret, C., Mirza, M. S., and Moënne-Loccoz, Y. (2006). Phylogeny of the 1-aminocyclopropane-1-carboxylic acid deaminase-encoding gene acdS in phytobeneficial and pathogenic Proteobacteria and relation with strain biogeography. *FEMS Microbiol. Ecol.* 56 455–470. doi: 10.1111/j.1574-6941.2006.00082.x

Carelli, M., Gnocchi, S., Fancelli, S., Mengoni, A., Paffetti, D., Scotti, C., et al. (2000). Genetic diversity and dynamics of *Sinorhizobium meliloti* populations nodulating different alfalfa varieties in Italian soils. *Appl. Environ. Microbiol.* 66, 4785–4789. doi: 10.1128/AEM.66.11.4785-4789.2000

Checcucci, A., Azzarello, E., Bazzicalupo, M., Galardini, M., Lagomarsino, A., Mancuso, S., et al. (2016). Mixed nodule infection in *Sinorhizobium meliloti* - *Medicago sativa* symbiosis suggest the presence of cheating behavior. *Front. Plant Sci.*. 7:835. doi: 10.3389/fpls.2016.00835

Crook, M. B., Lindsay, D. P., Biggs, M. B., Bentley, J. S., Price, J. C., Clement, S. C., et al. (2012). Rhizobial plasmids that cause impaired symbiotic nitrogen fixation and enhanced host invasion. *Mol. Plant Microbe Interact.* 25, 1026–1033. doi: 10.1094/mpmi-02-12-0052-r

Deikman, T. (1997). Molecular mechanisms of ethylene regulation of gene transcription. *Physiol. Plant.* 100, 561–566. doi: 10.1111/j.1399-3054.1997.tb03061.x

diCenzo, G. C., MacLean, A. M., Milunovic, B., Golding, G. B., and Finan, T. (2014). Examination of prokaryotic multipartite genome evolution through experimental genome reduction. *PLoS Genet.* 10:e1004742. doi: 10.1371/journal.pgen.1004742

Duan, J., Müller, K. M., Charles, T. C., Vesely, S., and Glick, B. R. (2009). 1-Aminocyclopropane-1-Carboxylate (ACC) deaminase genes in rhizobia from Southern Saskatchewan. *Microb. Ecol.* 57, 423–436. doi: 10.1007/s00248-008-9407-6

Ehrhardt, D. W., Atkinson, E. M., and Long, S. R. (1992). Depolarization of alfalfa root hair membrane potential by *Rhizobium meliloti* Nod factors. *Science* 256, 998–1000. doi: 10.1126/science.10744524

Galardini, M., Brilli, M., Spini, G., Rossi, M., Roncaglia, B., Bani, A., et al. (2015). Evolution of intra-specific regulatory networks in a multipartite bacterial genome. *PLoS Comput. Biol.* 11:e1004478. doi: 10.1371/journal.pcbi.1004478

Galardini, M., Mengoni, A., Brilli, M., Pini, F., Fioravanti, A., Lucas, S., et al. (2011). Exploring the symbiotic pangenome of the nitrogen-fixing bacterium *Sinorhizobium meliloti*. *BMC Genomics* 12:235. doi: 10.1186/1471-2164-12-235

Galardini, M., Pini, F., Bazzicalupo, M., Biondi, E. G., and Mengoni, A. (2013). Replicon-dependent bacterial genome evolution: the case of *Sinorhizobium meliloti*. *Genome Biol. Evol.* 5, 542–558. doi: 10.1093/gbe/evt027

Gamalero, E., and Glick, B. R. (2015). Bacterial modulation of plant ethylene levels. *Plant Physiol.* 169, 13–22. doi: 10.1104/pp.15.00284

Glick, B. R. (2005). Modulation of plant ethylene levels by the bacterial enzyme ACC deaminase. *FEMS Microbiol. Lett.* 251, 1–7. doi: 10.1016/j.femsle.2005.07.030

Glick, B. R., Penrose, D. M., and Li, J. (1998). A model for the lowering of plant ethylene concentrations by plant growth-promoting bacteria. *J. Theor. Biol.* 190, 63–68. doi: 10.1006/jtbi.1997.0532

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., et al. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* 38, W695–W699. doi: 10.1093/nar/gkq313

Grichko, V. P., and Glick, B. R. (2000). Identification of DNA sequences that regulate the expression of the Enterobacter cloacae UW4 1-aminocyclopropane-1-carboxylic acid deaminase gene. *Can. J. Microbiol.* 46, 1159–1165. doi: 10.1139/cjm-46-12-1159

Hammer, Ø., Harper, D. A. T., and Ryan, P. D. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* 41:9. Available online at: http://palaeo-electronica.org/2001_1/past/issue1_01.htm

Holguin, G., and Glick, B. R. (2003). Transformation of *Azospirillum brasilense* Cd with an ACC deaminase gene from *Enterobacter cloacae* UW4 fused to the Tet r gene promoter improves its fitness and plant growth promoting ability. *Microb. Ecol.* 46, 122–133. doi: 10.1007/s00248-002-1036-x

Honma, M., and Shimomura, T. (1978). Metabolism of 1-aminocyclopropane-1-carboxylic acid. *Agric. Biol. Chem.* 42, 1825–1831. doi: 10.1271/bbb1961.42.1825

Hontzeas, N., Richardson, A. O., Belimov, A., Safronova, V., Abu-Omar, M. M., and Glick, B. R. (2005). Evidence for horizontal transfer of 1-aminocyclopropane-1-carboxylate deaminase genes. *Appl. Environ. Microbiol.* 71, 7556–7558. doi: 10.1128/AEM.71.11.7556-7558.2005

Karunakaran, R., Mauchline, T. H., Hosie, A. H. F., and Poole, P. S. (2005). A family of promoter probe vectors incorporating autofluorescent and chromogenic reporter proteins for studying gene expression in Gram-negative bacteria. *Microbiology* 151, 3249–3256. doi: 10.1099/mic.0.28311-0

Khan, S. R., Gaines, J., Roop, R. M., and Farrand, S. K. (2008). Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. *Appl. Environ. Microbiol.* 74, 5053–5062. doi: 10.1128/AEM.01098-08

Kong, Z., Mohamad, O. A., Deng, Z., Liu, X., Glick, B. R., and Wei, G. (2015). Rhizobial symbiosis effect on the growth, metal uptake, and antioxidant responses of *Medicago lupulina* under copper stress. *Environ. Sci. Pollut. Res.* 22, 12479–12489. doi: 10.1007/s11356-015-4530-7

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Lemaire, B., Van Cauwenberghe, J., Chimphango, S., Stirton, C., Honnay, O., Smets, E., et al. (2015). Recombination and horizontal transfer of nodulation and ACC deaminase (*acdS*) genes within *Alpha-* and *Betaproteobacteria* nodulating legumes of the Cape Fynbos biome. *FEMS Microbiol. Ecol.* 91:fiv118. doi: 10.1093/femsec/fiv118

Li, J., and Glick, B. R. (2001). Transcriptional regulation of the *Enterobacter cloaceae* UW4 1-aminocyclopropane-1-carboxylate (ACC) deaminase gene (AcdS). *Can. J. Microbiol.* 47, 259–267. doi: 10.1139/cjm-47-4-359

Lugtenberg, B. J. J., Chin-A-Woeng, T. F. C., and Bloemberg, G. V. (2002). Microbe-plant interactions: principles and mechanisms. *Antonie Van Leeuwenhoek* 81, 373–383. doi: 10.1023/A:1020596903142

Ma, W., Charles, T. C., and Glick, B. R. (2004). Expression of an exogenous 1-aminocyclopropane-1-carboxylate deaminase gene in *Sinorhizobium meliloti* increases its ability to nodulate alfalfa. *Appl. Environ. Microbiol.* 70, 5891–5897. doi: 10.1128/AEM.70.10.5891-5897.2004

Ma, W., Guinel, F. C., and Glick, B. R. (2003a). *Rhizobium leguminosarum* biovar viciae 1-aminocyclopropane-1-carboxylate deaminase promotes nodulation of pea plants. *Appl. Environ. Microbiol.* 69, 4396–4402. doi: 10.1128/AEM.69.8.4396-4402.2003

Ma, W., Sebestianova, S. B., Sebestian, J., Burd, G. I., Guinel, F. C., and Glick, B. R. (2003b). Prevalence of 1-aminocyclopropane-1-carboxylate deaminase in Rhizobium spp. *Antonie Van Leeuwenhoek* 83, 285–291. doi: 10.1023/A:1023360919140

Markowitz, V. M., Chen, I. M., Palaniappan, K., Chu, K., Szeto, E., Pillay, M., et al. (2013). IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res.* 42, D560–D567. doi: 10.1093/nar/gkt963

Meade, H. M., Long, S. R., Ruvkun, G. B., Brown, S. E., and Ausubel, F. M. (1982). Physical and genetic characterization of symbiotic and auxotrophic mutants of

Rhizobium meliloti induced by transposon Tn5 mutagenesis. *J. Bacteriol.* 149, 114–122.

Mursel, V., Hennecke, H., and Pessi, G. (2012). Disparate role of rhizobial ACC deaminase in root-nodule symbioses. *Symbiosis* 57, 43–50. doi: 10.1007/s13199-012-0177-z

Nascimento, F. X., Brígido, C., Glick, B. R., and Oliveira, S. (2012). ACC deaminase genes are conserved among Mesorhizobium species able to nodulate the same host plant. *FEMS Microbiol. Lett.* 336, 26–37. doi: 10.1111/j.1574-6968.2012.02648.x

Nascimento, F. X., Rossi, M. J., Soares, C. R. F. S., McConkey, B. J., and Glick, B. R. (2014). New insights into 1-Aminocyclopropane-1-Carboxylate (ACC) deaminase phylogeny, evolution and ecological significance. *PLoS ONE* 9:e99168. doi: 10.1371/journal.pone.0099168

Nei, M., and Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. New York, NY: Oxford University Press.

Nukui, N., Minamisawa, K., Ayabe, S-I., and Aoki, T. (2006). Expression of the 1-aminocyclopropane-1-carboxylic acid deaminase gene requires symbiotic nitrogen-fixing regulator gene nifA2 in *Mesorhizobium loti* MAFF303099. *Appl. Environ. Microbiol.* 72, 4964–4969. doi: 10.1128/AEM.02745-05

Ogawa, J., and Long, S. R. (1995). The *Rhizobium meliloti* groELc locus is required for regulation of early nod genes by the transcription activator NodD. *Genes Dev.* 9, 714–729. doi: 10.1101/gad.9.6.714

Penrose, D. M., and Glick, B. R. (2003). Methods for isolating and characterizing ACC deaminase-containing plant growth-promoting rhizobacteria. *Physiol. Plant.* 118, 10–15. doi: 10.1034/j.1399-3054.2003.00086.x

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

Pini, F., Frage, B., Ferri, L., De Nisco, N. J., Mohapatra, S. S., Taddei, L., et al. and Biondi, E.G. (2013). The DivJ, CbrA and PleC system controls DivK phosphorylation and symbiosis in *Sinorhizobium meliloti*. *Mol. Microbiol.* 90, 54–71. doi: 10.1111/mmi.12347

Pini, F., Galardini, M., Bazzicalupo, M., and Mengoni, A. (2011). Plant-bacteria association and symbiosis: are there common genomic traits in Alphaproteobacteria? *Genes* 2, 1017–1032. doi: 10.3390/genes2041017

Pini, F., Spini, G., Galardini, M., Bazzicalupo, M., Benedetti, A., Chiancianesi, M., et al. (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. doi: 10.1007/s11104-013-1979-3

Prigent-Combaret, C., Blaha, D., Pothier, J. F., Vial, L., Poirier, M. A., Wisniewski-Dyé, F., et al. (2008). Physical organization and phylogenetic analysis of acdR as leucine-responsive regulator of the 1-aminocyclopropane-1-carboxylate deaminase gene acdS in phytobeneficial *Azospirillum lipoferum* 4B and other Proteobacteria. *FEMS Microbiol. Ecol.* 65, 202–219. doi: 10.1111/j.1574-6941.2008.00474.x

Reid, J. B., and Renquist, A. R. (1997). Enhanced root production as a feed-forward response to soil water deficit in field-grown tomatoes. *Funct. Plant Biol.* 24, 685–692. doi: 10.1071/pp96079

Roumiantseva, M. L., Andronov, E. E., Sharypova, L. A., Dammann-Kalinowski, T., Keller, M., Young, J. P. W., et al. (2002). Diversity of *Sinorhizobium meliloti* from the central Asian alfalfa gene center. *Appl. Environ. Microbiol.* 68, 4694–4697. doi: 10.1128/aem.68.9.4694-4697.2002

Roumiantseva, M. L., Muntyan, V. S., Mengoni, A., and Simarov, B. V. (2014). ITS-polymorphism of salt-tolerant and salt-sensitive native isolates of *Sinorhizobium meliloti*-symbionts of alfalfa, clover and fenugreek plants. *Russ. J. Genet.* 50, 348–359. doi: 10.1134/s1022795414040103

Simon, R., Priefer, U., and Pühler, A. (1983). A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram negative bacteria. *Nat. Biotechnol.* 1, 784–791. doi: 10.1038/nbt1183-784

Schneiker-Bekel, S., Wibberg, D., Bekel, T., Blom, J., Linke, B., Neuweger, H., et al. (2011). The complete genome sequence of the dominant *Sinorhizobium meliloti* field isolate SM11 extends the *S. meliloti* pan-genome. *J. Biotechnol.* 155, 20–33. doi: 10.1016/j.jbiotec.2010.12.018

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464. doi: 10.1214/aos/1176344136

Singh, R. P., Shelke, G. M., Kumar, A., and Jha, P. N. (2015). Biochemistry and genetics of ACC deaminase: a weapon to "stress ethylene" produced in plants. *Front. Microbiol.* 6:937. doi: 10.3389/fmich.2015.00937

Skouloubris, S., Labigne, A., and De Reuse, H. (1997). Identification and characterization of an aliphatic amidase in Helicobacter pylori. *Mol. Microbiol.* 25, 989–998.

Stiens, M., Schneiker, S., Keller, M., Kuhn, S., Pühler, A., and 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. doi: 10.1128/aem.72.5.3662-3672.2006

Sun, Y., Cheng, Z., and Glick, B. R. (2009). The presence of a 1-aminocyclopropane-1-carboxylate (ACC) deaminase deletion mutation alters the physiology of theendophytic plant growth-promoting bacterium *Burkholderia phytofirmans*. *FEMS Microbiol Lett.* 296, 131–136. doi: 10.1111/j.1574-6968.2009.01625.x

Talebi Bedaf, M. B., Bahar, M., Saeidi, G., Mengoni, A., and Bazzicalupo, M. (2008). Diversity of Sinorhizobium strains nodulating *Medicago sativa* from different Iranian regions. *FEMS Microbiol. Lett.* 288, 40–46. doi: 10.1111/j.1574-6968.2008.01329.x

Trabelsi, D., Mengoni, A., Aouani, M. E., Bazzicalupo, M., and Mhamdi, R. (2010). Genetic diversity and salt tolerance of *Sinorhizobium* populations from two Tunisian soils. *Ann. Microbiol.* 60, 541–547. doi: 10.1007/s13213-010-0084-6

Trabelsi, D., Pini, F., Aouani, M. E., Bazzicalupo, M., and Mengoni, A. (2009). Development of real-time PCR assay for detection and quantification of *Sinorhizobium meliloti* in soil and plant tissue. *Lett. Appl. Microbiol.* 48, 355–361. doi: 10.1111/j.1472-765X.2008.02532.x

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

Chapter 4

4.1 Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*

The multipartite genome organization of a large variety of bacteria, and in particular of our model species *Sinorhizobium meliloti*, kicked off the questions about its origin and functional meaning in evolutionary terms. In particular, the genomic composition of the replicons allowed to speculate a different role of each one in the microbial life cycle and possibly, an adaptative function in different environmental conditions.

nature COMMUNICATIONS

ARTICLE

Received 11 Dec 2015 | Accepted 10 Jun 2016 | Published 22 Jul 2016

DOI: 10.1038/ncomms12219 OPEN

Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*

George C. diCenzo[1], Alice Checcucci[2], Marco Bazzicalupo[2], Alessio Mengoni[2], Carlo Viti[3], Lukasz Dziewit[4], Turlough M. Finan[1], Marco Galardini[5] & Marco Fondi[2]

The genome of about 10% of bacterial species is divided among two or more large chromosome-sized replicons. The contribution of each replicon to the microbial life cycle (for example, environmental adaptations and/or niche switching) remains unclear. Here we report a genome-scale metabolic model of the legume symbiont *Sinorhizobium meliloti* that is integrated with carbon utilization data for 1,500 genes with 192 carbon substrates. Growth of *S. meliloti* is modelled in three ecological niches (bulk soil, rhizosphere and nodule) with a focus on the role of each of its three replicons. We observe clear metabolic differences during growth in the tested ecological niches and an overall reprogramming following niche switching. *In silico* examination of the inferred fitness of gene deletion mutants suggests that secondary replicons evolved to fulfil a specialized function, particularly host-associated niche adaptation. Thus, genes on secondary replicons might potentially be manipulated to promote or suppress host interactions for biotechnological purposes.

[1] Department of Biology, McMaster University, Hamilton, Ontario, Canada L8S 1A1. [2] Department of Biology, University of Florence, 50019 Sesto Fiorentino, Italy. [3] Department of Agri-food Production and Environmental Sciences, University of Florence, 50144 Sesto Fiorentino, Italy. [4] Department of Bacterial Genetics, Institute of Microbiology, Faculty of Biology, University of Warsaw, 02-096 Warsaw, Poland. [5] EMBL-EBI, Wellcome Trust Genome Campus, Cambridge CB10 1SD, UK. Correspondence and requests for materials should be addressed to A.M. (email: alessio.mengoni@unifi.it).

The last years have witnessed a growing attention towards the ecological and evolutionary implication of the multiple replicon bacterial genome[1–4] that is present in about 10% of sequenced bacterial genomes[1]. This genome architecture is common in the proteobacterial species that interact with a host and are of importance to the human population[1,2], including crop plant symbionts (for example, *Sinorhizobium* and *Rhizobium*), plant pathogens (for example, *Agrobacterium*), and animal and human pathogens (for example, *Brucella*, *Burkholderia* and *Vibrio*). As the bacterial genome is non-randomly organized[5], it is proposed that this genome organization was shaped by selective pressures to facilitate improved host interactions and niche adaptation. Though it is well established that secondary replicons often carry genetic determinants essential to colonize a novel environment, for example, virulence or symbiotic genes, such genes often account for only a small proportion of these replicons[6,7]. The majority of the genes on a secondary replicon are not directly essential to colonize a specific environment, and the adaptive function of these genes and why they are localized on a secondary replicon remains unclear. Several recent studies have provided evidence consistent with the secondary replicons in a multipartite genome encoding environment-specific fitness promoting but non-essential functions[3,4,8–12]. However, none of these studies demonstrated that secondary replicons indeed carry environment-specific fitness determinants, thus serving as reservoirs for niche-specific functions.

Sinorhizobium meliloti is a N_2-fixing endosymbiont of legume species that has recently become a model organism for the study of bacterial multipartite genome function and evolution. All sequenced *S. meliloti* genomes contain at least three large replicons (the primary chromosome, the pSymB chromid and the pSymA megaplasmid), with some strains hosting additional small accessory plasmids[3,13,14]. *S. meliloti* experiences a complex life cycle and successfully colonizes three distinct niches. Two of these niches are bulk soil and rhizosphere soil (that is, the soil directly influenced by the plant root system), which are quite different environments, with the rhizosphere generally considered to be a nutritionally richer environment due to plant root exudates[15]. The third niche inhabited by *S. meliloti* is the legume root nodule. *S. meliloti* can induce root nodule formation in certain legumes and within nodules the bacteria differentiate into N_2-fixing bacteroids. Manipulation and optimization of this agriculturally and ecologically important symbiosis is an ultimate goal of the rhizobial research community[16–18]. Effectively doing so will require a complete understanding of the evolution[18], genetics[6] and the metabolism of the organism in both rhizosphere and nodule environments, as well as the corresponding metabolic shifts.

Here we combine a genome-scale metabolic network reconstruction of the *S. meliloti* genome, flux balance analysis (FBA), and growth phenotype data for 11 large-scale *S. meliloti* deletion mutants to examine the metabolic changes accompanying the shifts between bulk soil, rhizosphere and nodule environments. We use an *in silico* approach to predict the phenotypes resulting from the deletion of 1,575 *S. meliloti* metabolic genes, estimate the fitness contribution of each replicon within each environment, and thus provide insight into the evolution of multipartite genomes.

Results

Reconstruction of a *S. meliloti* genome-scale metabolic model. As described in the Supplementary Methods, an *in silico* representation of the metabolism of *S. meliloti* was developed, and the final model that was termed iGD1575 contained 1,575 genes, 1,825 reactions and 1,579 metabolites. iGD1575 accounts for 25.4% of the protein-coding genes in the *S. meliloti* genome, and the other main features of the model are listed in Table 1. Cluster of Orthologous Gene (COG) analyses confirmed that the gene functional biases of each replicon are accurately represented in iGD1575 (Supplementary Fig. 1)[12,13]. The iGD1575 model encompasses 529 of the 565 genes present in iHZ565, a previously reported *S. meliloti* small metabolic model[19]. The remaining 31 genes were not added to iGD1575 as experimental data were inconsistent with their annotation, we felt their annotation too general to have high confidence in the enzymes' substrates/products, or the associated reaction involved a metabolite not present in any other reaction in the model and thus the reaction would never carry flux in FBA (Supplementary Table 1). Comparison of the number of genes in iGD1575 to that of other available rhizobial and non-rhizobial models[19–22] showed that iGD1575 is currently one of the largest metabolic reconstruction of a bacterial genome. In addition, iGD1575 is the first metabolic model capable of representing the metabolism of both a symbiotic and free-living rhizobial cell.

Quantitative validation of iGD1575. Previous work[23] has shown that *S. meliloti* transports glucose into the cell at a rate of 2.41 mmol h^{-1} per g cellular dry weight. When glucose is provided to the iGD1575 model as the sole carbon source at the experimentally determined rate, a specific growth rate of 0.325 h^{-1} is predicted, which is consistent with our experimentally derived growth rate of 0.313 h^{-1} (s.d. 0.002) for *S. meliloti* grown with glucose. Similarly, it has been shown[24] that *S. meliloti* transports succinate into the cell at a rate of 6.252 mmol h^{-1} per g protein. Providing succinate as the sole source of carbon to iGD1575 at the experimentally determined value led to a predicted specific growth rate of 0.279 h^{-1}, similar to the experimentally derived growth rate of 0.254 h^{-1} (s.d. 0.025) for *S. meliloti* grown with succinate. Measuring the amount of phosphate remaining in the spent growth media following growth of *S. meliloti* indicated that 63.7 (s.d. 6.7) and 39.8 μM (s.d. 1.5) of phosphate was used per mM of glucose and succinate, respectively. These experimental values are relatively consistent with the phosphate usage values predicted by iGD1575 of 72.7 and 48.5 μM per mM of glucose and succinate, respectively.

Little experimental flux data has been reported for *S. meliloti*; however, flux measurements for 22 central carbon metabolic reactions when *S. meliloti* is grown with glucose as the sole source of carbon have been reported by Fuhrer *et al.*[23]. Not surprisingly, the experimentally determined fluxes did not match well with the iGD1575-derived values. This is because the specific growth rate of *S. meliloti* was only 0.17 h^{-1} in the study by Fuhrer *et al.*, indicating that *S. meliloti* was grown in sub-optimal conditions that presumably effected the flux distribution. Nevertheless, if the flux through these 22 central carbon metabolic reactions in iGD1575 was set to that as experimentally determined by Fuhrer *et al.*, the predicted specific growth rate was reduced to 0.159 h^{-1}, in line with the 0.17 h^{-1} reported by Fuhrer *et al.* The good relationship between flux distribution and specific growth rate, and the strong ability of iGD1575 to predict growth rate and phosphate usage when grown with glucose or succinate, suggest that the flux distributions predicted by iGD1575 should represent quantitatively accurate estimations.

iGD1575 captures the metabolic capacity of *S. meliloti*. The ability of *S. meliloti* to grow with various carbon and nitrogen sources has been well studied by means of the Phenotype MicroArray (Biolog) technology[4,25–27]. These previously published studies were used to guide model expansion and refinement during the curation process. Once all of the manual curation of iGD1575 was complete, FBA illustrated that the final model could accurately

NATURE COMMUNICATIONS | DOI: 10.1038/ncomms12219

Table 1 | Summary of the main properties of iGD1575.

S. meliloti 1021 genome	
Total genome size	6,691,694
Size of the chromosome (% total)	3,654,135 (54.6)
Size of pSymB (% total)	1,683,333 (25.2)
Size of pSymA (% total)	1,354,226 (20.2)
Total protein-coding genes (PCG)	6 204
Chromosome PCG (% total)	3,341 (53.9)
pSymB PCG (% total)	1,570 (25.3)
pSymA PCG (% total)	1,293 (20.8)
iGD1575 characteristics	
Total genes (% of *S. meliloti* genes)	1,575 (25.4)
Chromosome genes (% total)	944 (69.9)
pSymB genes (% total)	390 (24.8)
pSymA genes (% total)	241 (15.3)
Total reactions (rxns)	1,825
Gene-associated rxns (gar) (% total)	1,404 (76.9)
Chromosome dependent (% gar)	898 (64.0)
pSymB dependent (% gar)	205 (14.6)
pSymA dependent (% gar)	73 (5.2)
Multiple replicons (% gar)	228 (16.2)
Unknown metabolic GPR (% total)	63 (3.4)
Unknown transport GPR (% total)	46 (2.5)
Exchange reactions (% total)*	270 (14.8)
Demand reactions (% total)†	22 (1.2)
Diffusion reactions (% total)	8 (0.4)
Spontaneous reactions (% total)	10 (0.5)
Objective functions (% total)	3 (0.2)
Total metabolites	1,579

GPR, Gene-Protein-Reaction.
*Exchange reactions are used to define the medium components.
†Demand reactions are used to provide compounds whose synthesis is not represented in the model. In all, 20 of the demand reactions represent the uncharged tRNA molecules. 2 are for fatty acids.

predict the ability of *S. meliloti* to produce, or not produce, biomass (as defined in Supplementary Table 2) on 85% (138/162) and 75% (64/85) of the tested carbon and nitrogen substrates, respectively, for which the ability of *S. meliloti* to utilize, or not, these compounds is known (Fig. 1 and Supplementary Data 1). Most of the discrepancies between the experimental data and the iGD1575 growth prediction were false negatives (71% and 95% for growth with carbon and nitrogen substrates, respectively). These represent compounds that *S. meliloti* can metabolize but the model cannot use for the production of biomass, likely representing gene annotation gaps in our knowledge of *S. meliloti* that will serve as targets for future research. The predictive power of bacterial metabolic models reported in previous studies[28–30] is similar to that reported here for iGD1575. Hence, iGD1575 is at least as good as other current genome-scale metabolic reconstructions at representing the organism's metabolic capabilities. This suggests that iGD1575 effectively captures the metabolic capacity of *S. meliloti* and can validly be used to model metabolism in nutritionally diverse environments.

Carbon growth phenotypes of *S. meliloti* deletion mutants. Carbon utilization phenotypes for a subset of large-scale pSymB deletion mutants[31] that cumulatively remove ~1.65 Mb (98%) of pSymB (Supplementary Fig. 2) were determined using PM1 and PM2A Biolog plates. This screen effectively generated a carbon utilization data set for ~1,500 pSymB genes. Overall, growth was observed with 76 carbon substrates, and a total of 43 no or extremely poor growth phenotypes were observed (Table 2, Supplementary Figs 3 and 4, and Supplementary Data 1 and 2). In the process of developing and validating iGD1575, an *in silico* representation of the same experiment was performed, and

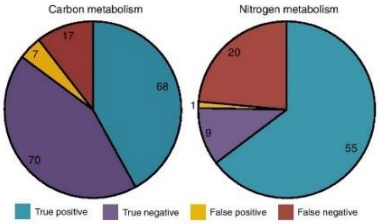


Figure 1 | Agreement between experimental and *in silico* metabolic capabilities of *S. meliloti*. True positives, growth was observed experimentally and *in silico*. True negatives, growth was not observed experimentally or *in silico*. False negatives, compounds that support growth experimentally but not *in silico*. False positives, compounds that support growth *in silico* but not experimentally. The complete set of compounds and growth predictions can be found in Supplementary Data 1.

where possible, the model was updated to fix discrepancies between the experimental and *in silico* results. Following this integration of the Phenotype MicroArray data set with the metabolic reconstruction, there was very good agreement between the experimentally observed results and the *in silico* simulations was observed (Table 2 and Supplementary Data 1). *In silico* simulations did not predict any 'no growth' phenotypes that were not experimentally observed, and 23 of the 36 (63.9%) experimentally observed phenotypes for compounds that support growth of iGD1575 were replicated *in silico*. Some of the discrepancies between the experimental and *in silico* data represent gaps in our knowledge of catabolic pathways in *S. meliloti*, while other phenotypes may occur for non-metabolic reasons and therefore not give a phenotype *in silico*. For example, the *S. meliloti* deletion mutant ΔB154 is more sensitive to cobalt chelation than the wild type[32], and the lack of growth in wells with L-histidine or D-glucosamine may simply reflect cobalt chelation[33,34].

In addition to model refinement, integrating the mutant phenotype data with iGD1575, the DuctApe software[27], the *S. meliloti* genome annotation[13] and an ABC transporter induction study[35] allowed for the prediction of novel carbon catabolic loci. One example compound is the monosaccharide psicose. Our analysis suggests that psicose is transported by the SupABCD (Smb20484–Smb20487) ABC transporter, and then converted to fructose by an isomerase encoded by *smb20488*. A second example is D-galactosamine, which, as elaborated on in Supplementary Note 1, we predict is transported by the Smb21216, Smb21219–Smb21221 transporter and potentially the Smb21135–Smb21138 transporter, and then metabolized by Smb21217, Smb21218, Smb21373 and Smb21374.

Rhizosphere colonization required a metabolic refinement. The metabolic shifts experienced by *S. meliloti* during transition between bulk soil, the rhizosphere and the nodule were modelled using *in silico* representations of the nutritional composition of each environment. These took into account the relative ratios of each component in the different environments and the development of these environments are described in the Supplementary Methods. In the bulk soil and rhizosphere environments, the model was optimized for the production of biomass as defined in Supplementary Table 2, whereas in the nodule environment the model was optimized for production of

Table 2 | Carbon utilization phenotypes observed for pSymB deletion mutants.

Strain	Phenotypes		
	Biolog	**Both**	**Model**
ΔB154	L-histidine D-glucosamine	None	None
ΔB141	Palatinose* Maltitol‡	D-trehalose *Trans*-4-hydroxy-L-proline	*Cis*-4-hydroxy-D-proline†
ΔB163	Dulcitol* D-glucosamine	D-psicose	None
ΔB180	Uridine* *N*-acetyl-D-galactosamine* Arbutin* D-raffinose Organic acids‖	L-leucine	D-galactosamine§
ΔB181	L-lysine* Dulcitol* *N*-acetyl-D-galactosamine* D-glucosamine	L-histidine D-tagatose	D-galactosamine§
ΔB108	L-ornithine L-serine L-asparagine L-alanine	None	None
ΔB109	None	None	None
ΔB179	Glycerol L-lactic acid	L-arabinose M-inositol	None
ΔB118	None	L-ornithine M-inositol L-arginine	None
ΔB182	Acetic acid Asparagine‡ α-methyl-D-galactoside‡ Melibionic acid‡ Bromosuccinate‡	D-melibiose L-malic acid D-raffinose Succinic acid D,L-malate Fumaric acid L-aspartate	Protocatechuate†
ΔB161	Dulcitol* α-methyl-D-galactoside‡ Melibionic acid‡	D-melibiose D-raffinose α-D-lactose Lactulose Methyl-β-D-galactoside	None

The 'Biolog' column indicates the phenotypes observed experimentally that are not seen *in silico* with the iGD1575 model, and *vice versa* for the 'Model' column. The 'Both' column lists the phenotypes observed both experimentally and *in silico*.
*The model does not produce biomass with this compound.
†This compound is not present in the PM1 or PM2A plates, but the phenotype is confirmed in the literature.
‡The model does not include this compound.
§This compound is not present in the PM1 or PM2A plates, and the phenotype has not been reported in the literature.
‖Includes all tested L-amino acids, gly-glu, ala-gly, gly-asp, L-lactic acid, acetic acid, pyruvic acid, methylpyruvic acid, melibionic acid and gamma-aminobutyric acid (GABA).

an effective N_2-fixing symbiosis as defined previously[19]. The optimal flux patterns in each of the three niches were obtained using FBA and visualized with iPath (Fig. 2)[36].

The metabolic network appears globally similar in both the bulk soil and rhizosphere environments (Fig. 2 and Table 3), although many subtle differences were present when reaction specific parameters were examined (Fig. 2 and Supplementary Data 3). Despite good correlation between the $\log_{10}$ of the absolute flux through a given reaction that was active in both environments (P value <0.01 using a Spearman's Rank Order Correlation test, median (absolute residual/observed) $= 0.09$; Supplementary Fig. 5a), $\sim$20% of the reactions showed at least a 50% change in flux between the two environments while an additional 6% switched directions. Similarly, the effect on fitness (defined as the flux through the objective function (biomass formation or symbiosis) in the mutant relative to the flux through the objective function in the wild type) of individual reaction deletions displayed a strong correlation between the two environments (P value <0.01 using a Spearman's Rank Order Correlation test, $R^2 = 0.95$; Supplementary Fig. 5b). Nevertheless, $\sim$7% had at least a 10% variation in fitness effect between environments, and $\sim$4% were essential in just one of the two niches. Interestingly, optimal growth in the rhizosphere required a greater repertoire of metabolic reaction as illustrated by the increased number of reactions required for maximal fitness. In addition, $\sim$13% of the active reactions were specific to just one of the environments. The reactions whose fluxes were considered to change between growth in bulk soil and the rhizosphere were further validated through a procedure involving flux variability analysis as detailed in Supplementary Note 2.

Few outstanding biases (P value <0.01 using a Pearson's χ^2-test) were seen in the COG annotations of the genes associated with reactions whose flux or fitness contribution was dependent on the soil environment. This indicated that the reactions important in the rhizosphere were biologically similar to, but functionally distinct from, the reactions important in bulk soil. However, coenzyme transport and metabolism (COG H), and cell wall, membrane and envelope biogenesis (COG M) were more important in the rhizosphere than in the bulk soil. This possibly reflects different coenzyme requirements for the metabolic

NATURE COMMUNICATIONS | DOI: 10.1038/ncomms12219 ARTICLE

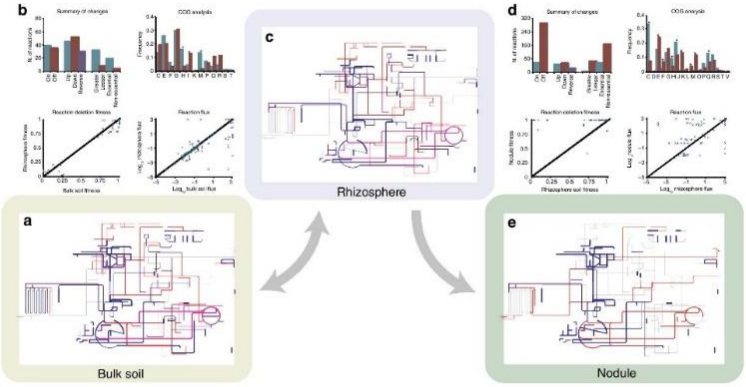


Figure 2 | The effect of niche conditions on the reconstructed metabolic network. Networks were visualized following optimization in (**a**) bulk soil, (**c**) rhizosphere and (**e**) nodule environments. Lines are colour coded based on fitness effect of deleting each reaction: blue indicates a fitness decrease <1%; dark purple indicates a fitness decrease <50%; bright purple indicates a fitness decrease >50%; and red indicates a fitness decrease >99%. Thin grey lines indicate inactive reactions. Line thickness shows the flux through each reaction on a log scale. The graphs summarize the metabolic changes detected during the (**b**) bulk soil to rhizosphere and (**d**) rhizosphere to nodule transitions. Summary of changes graphs: on and off—reactions carrying flux only in the second and first environment, respectively; up and down—reactions carrying increased flux (≥50%) in the second and first environment, respectively; reverse—reactions whose directionality is switched; greater and lesser—reactions whose removal have a greater (≥10%) fitness impact in the second and first environment, respectively; essential and non-essential—reactions essential only in the second and first environment, respectively. The nine classifications are not mutually exclusive. The reactions present in each category are described in Supplementary Data 3 and 4. The COG analysis graphs summarize the functional annotation of the genes associated with the reactions in the summary of changes graphs. The blue and red bars include the genes associated with the blue and red bars, respectively, in the summary of changes graphs. Asterisks indicate statistically significant changes (*P* value <0.01) as determined by Pearson's χ^2-tests. In the reaction flux figures, each point represents the amount of flux through individual reactions in the two environments. Blue and purple symbols indicate reactions with the same or reverse directionality, respectively. The angled line indicates the position of a perfect correlation. In the reaction deletion fitness figures, each point represents the fitness of individual reaction deletion mutants in the two environments. The angled line indicates the position of a perfect correlation.

pathways active in the two environments and the increased succinoglycan content of *S. meliloti* in the rhizosphere that is necessary to facilitate root biofilm formation. Lipid transport and metabolism (COG I) was over-represented in the bulk soil, perhaps due to the over-abundance of ketogenic amino acids in bulk soil. At the pathway level, only a few changes could not be explained by differences in the nutritional composition and biomass objective functions (Supplementary Data 3). For example, the importance of various carbon catabolic pathways and amino-acid biosynthetic pathways reflected the abundance of the sugars and amino acids in each environment. This analysis also revealed that *S. meliloti* relies more heavily on glycolytic substrate during growth in bulk soil but on gluconeogenic substrate in the rhizosphere, which was consistent with the high concentration of organic acids in the rhizosphere. The increased gluconeogenic flux and the increased flux through the pantothenate and coenzyme-A biosynthesis pathways in the rhizosphere is also consistent with an increased sugar demand for the rhizosphere-specific Nod factor production and increased exopolysaccharide biosynthesis[37,38]. Finally, the urea cycle contributed more to cellular fitness in bulk soil than in the rhizosphere.

Complex metabolic reprogramming is associated with symbiosis. The rhizosphere to nodule transition was accompanied with much more pronounced metabolic changes than the bulk soil to rhizosphere transition (Fig. 2 and Supplementary Data 4). Half as many reactions carried flux in the nodule than in the rhizosphere, with ~61% of rhizosphere reactions off in the nodule and ~22% of active nodule reactions off in the rhizosphere. This overall decrease in metabolic reactions active in the nodule is consistent with the global transcriptional downregulation in differentiated bacteroids[39,40]. For reactions active in both environments, there was a significant correlation (*P* value<0.01 using a Spearman's Rank Order Correlation test; Supplementary Fig. 5c) in the $\log_{10}$ of the absolute flux values, but the dispersion of the observed values from the regression line was high (median (residual/observed) = 1.48). Approximately half of the common flux carrying reactions displayed at least 50% more flux in one of the environments and a further 12% switched directions. In addition, little correlation was observed between the fitness effects of individual reaction deletions in the two environments ($R^2 = 0.03$; Supplementary Fig. 5d). Of the active reactions, ~38% were essential specifically in one environment, while the deletion of another 12% gave fitness effects ≥10% different in the two niches. The reactions whose fluxes were considered to change between growth in the rhizosphere and symbiosis in the nodule were further validated through a procedure involving flux variability analysis as detailed in Supplementary Note 2.

Table 3 | Nutritional composition of the rhizosphere and bulk soil.

Compound	Rhizosphere*	Bulk soil*
Total sugars	0.615	1.753
Arabinose	0.201	0.273
Fucose†	0.030	0.064
Galactose	0.136	0.239
Galacturonic acid‡	0.053	—
Glucose	0.030	0.586
Glucuronic acid‡	0.023	—
Mannose	0.040	0.240
Rhamnose	0.022	0.132
Xylose	0.034	0.201
Ribose	—	0.018
Sucrose	0.015	0.000
Raffinose	0.008	0.000
Stachyose	0.023	0.000
Total organic acids	1.231	0.072
Succinate	0.091	0.005
Malate	0.591	0.035
Citrate†	0.549	0.032
Total amino acids	0.154	0.175
Aspartic acid	0.024	0.000
Threonine	0.007	0.006
Serine	0.007	0.000
Homoserine†	0.046	—
Glutamic acid	0.021	0.000
Proline	0.000	0.010
Glycine	0.005	0.006
Alanine	0.006	0.026
Valine	0.003	0.028
Cysteine	0.000	0.006
Isoleucine	0.001	0.019
Leucine	0.002	0.034
Tyrosine	0.002	0.006
Phenylalanine‡	0.002	0.006
GABA	0.000	—
Ornithine	0.006	0.006
Lysine†	0.004	0.006
Histidine	0.004	0.006
Arginine	0.007	0.006
Asparagine	—	0.000
Glutamine	—	0.000
Hydroxyproline	0.006	—
Ammonium	Excess	Excess
Nitrate	Excess	Excess
Sulfate	Excess	Excess
Phosphate	Excess	Excess

GABA, gamma-aminobutyric acid.
*Values represent the molar ratio of the compounds, with the sum of all compounds in each environment equalling 2. —, no information.
†Excluded from the *in silico* representation as the model fails to grow on these compounds.
‡Excluded from the *in silico* representation as *S. meliloti* does not grow with these compounds.

A clear shift in the functional annotation of genes associated with the variable reactions was observed. Functions associated with generating the large amount of energy required for nitrogen fixation displayed increased importance in the nodule: for example, energy production and conversion (COG C), and coenzyme transport and metabolism (COG H). On the other hand, the lack of growth of the differentiated bacteroids not surprisingly rendered biomass component biosynthesis (COGs E, F, I., M, I and J) less important. A few additional interesting observations were noted by looking at pathway level changes (Supplementary Data 4). The Kreb's cycle and AMP synthesis were increased, presumably to accommodate the high ATP demand of nitrogenase. Glycolysis was less important in the nodule, consistent with the lack of glycolysis-specific enzymes detected in the *S. meliloti* nodule proteome[41]. Flux through various pathways producing compounds (including steroids, glutathionine, vitamin B6 and haem) required for a successful symbiosis was observed, and in most cases these changes are supported by previously published proteomic, RNA-seq or induction studies[41–43]. Flux through the non-oxidative pentose phosphate pathway, which is poorly studied in *S. meliloti*[44], was also increased, consistent with the detection of two enzymes of this pathway in the *S. meliloti* nodule proteome[41] and the need for *S. meliloti* to synthesize sugars for biosynthesis[45].

***S. meliloti* replicons encode niche-specific metabolism.** We performed comprehensive, replicon-specific *in silico* single and double gene deletion analyses to determine the contribution of the three *S. meliloti* replicons to the overall fitness of *S. meliloti* in each of the tested environments (Supplementary Table 3, Fig. 3 and Supplementary Fig. 6). The use of a double gene deletion analysis was intended to account for functionally redundancy gene pairs that would mask phenotypes during the single gene deletion analysis[46,47]. As before, fitness was determined as the flux through the objective function of the mutant relative to the wild type, with the biomass formation (Supplementary Table 2) being the objective function during growth in bulk soil and the rhizosphere, and N_2-fixation[19] being the objective function in the nodule environment.

The mutant analyses revealed that the *S. meliloti* chromosome had a similar contribution to fitness in bulk soil and the rhizosphere; there was little change in the number of essential or fitness-contributing chromosomal genes in these two environments (Fig. 3 and Supplementary Table 3). However, there was a clear reduction in the importance of the chromosome during symbiosis in the nodule, consistent with microarray data showing that chromosomal genes are over-represented amongst the genes that have low expression levels in the symbiotic bacteria relative to the free-living form[41].

Similar to the chromosome and consistent with the global *S. meliloti* transcriptional downregulation in the nodule[39], pSymB contributed more or less only to the fitness of the free-living bacterium, with little role detected in the bacteroids (Fig. 3 and Supplementary Table 3). However, unlike the chromosome, pSymB showed a bias in importance between growth in bulk soil and the rhizosphere; the number of fitness-promoting genes was ~3.5-fold greater in the rhizosphere. Moreover, every pSymB gene that contributed to fitness in bulk soil had a greater fitness contribution in the rhizosphere. This rhizosphere bias was further amplified when considering the origin of fitness-promoting genes. Of the five pSymB genes contributing to fitness in bulk soil, four are involved in arabinose transport or catabolism[48]. All four of these genes have a chromosomal origin and were transferred to pSymB through an inter-replicon translocation event[49]. We therefore detected only a single gene (*smb20201*) contributing to fitness in bulk soil that originated on pSymB. Similarly, transcriptomics work with the pea symbiont, *Rhizobium leguminosarum*, indicated that one of its plasmids (pRL8) is over-represented in genes upregulated specifically in the pea rhizosphere[50]. However, with a few exceptions, the fitness contributions of the pLR8 upregulated genes in bulk soil versus the rhizosphere were not determined.

Even though these data clearly illustrated that the metabolic capabilities encoded by pSymB are either specific or more important for growth in the rhizosphere than bulk soil, we believe that the observed bias is an under-representation of the actual situation. The succinoglycan biosynthetic genes are classified as essential in both the bulk soil and the rhizosphere due to their inclusion in the biomass objective functions; however, they are

NATURE COMMUNICATIONS | DOI: 10.1038/ncomms12219 ARTICLE

not truly essential but likely have greater importance in the rhizosphere through facilitating biofilm formation on the legume root. Furthermore, a more complete formulation of the bulk soil and rhizosphere environment may exaggerate the bias. For example, protocatechuate was not included due to a lack of information on its abundance. However, recent work showed that protocatechuate metabolism improved fitness of *R. leguminosarum* in the rhizosphere[51], and 13 pSymB genes are involved in protocatechuate transport and metabolism[52,53].

In contrast with the other replicons, the pSymA megaplasmid contributed no fitness-promoting genes (Fig. 3 and Supplementary Table 3). No phenotypes were detected in bulk soil, while the 'essential' genes in the rhizosphere were due to the removal of Nod factor biosynthetic genes. In fact, Nod factor biosynthesis is not essential for growth but is required for the initiation of symbiosis. In the nodule, the essential genes that were identified were required for the synthesis and functioning of the nitrogenase enzyme. The lack of fitness-contributing pSymA genes in the nodule was somewhat surprising, although consistent with published data[6,54], perhaps suggesting that few genes outside of the core symbiotic machinery contribute to the nitrogen fixation process. Indeed, the large rearrangements in the structure of pSymA between wild-type *S. meliloti* nodule isolates[3] may reflect low selective constraints on the pSymA megaplasmid, and thus explain the low metabolic contribution and importance of pSymA even during the symbiotic interaction.

The biases observed for the importance of each replicon in the different environments were confirmed via random permutations, testing up to 1,000 different nutritional compositions as described in detail in Supplementary Note 3. These environments were created by randomly varying, at each iteration, the maximal allowable uptake of each nutrient with respect to the original value and by also randomly removing two nutrients from the environment. Despite some interesting biological insights being derived from this analysis (Supplementary Note 3 and Supplementary Figs 7 and 8), little variation was seen in the number of essential plus fitness-contributing genes on each replicon or in each environment (Supplementary Figs 7 and 8). The robustness of these results to environmental variations provides support for the validity of our conclusions and shows that the niche specialization is not unique to the specific environmental composition used throughout this study.

Finally, a comparison of genes differentially contributing to growth in each environment with a recent regulon analysis in *S. meliloti*[12] was not conclusive due to the low overlap of the data sets (Supplementary Table 4 and Supplementary Data 5; additional details in Supplementary Note 4). On the other hand, grouping these genes based on their pangenome classification[12] illustrated that nearly all fitness-contributing genes belonged to the core genome, a clear enrichment relative to the percentage of core genes in iGD1575 overall (Supplementary Table 5 and Supplementary Data 5; additional details in Supplementary Note 5).

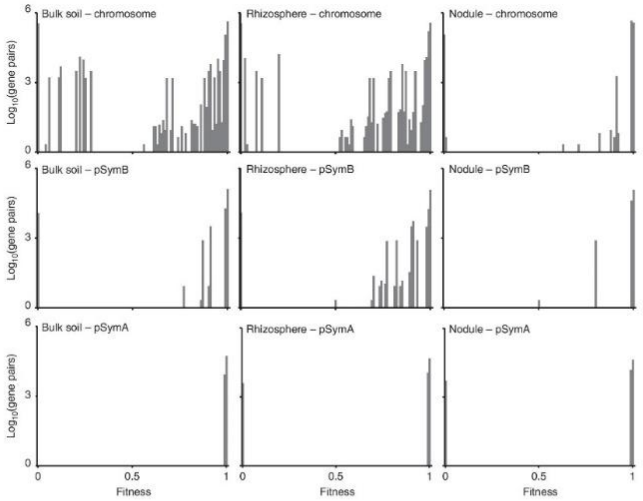


Figure 3 | Fitness costs associated with double gene deletions in the three tested ecological niches. All possible pairs of genes present on the same replicon were individually removed from the model, the ability of the resulting mutant to produce flux through the objective function was examined with FBA, and the fitness (solution value of the mutant/solution value of the wild type) of each mutant was calculated. The histograms summarize the calculated fitness values for each mutant in each of the three environments separately for each replicon. The fitness is displayed on the *x* axis, with the number of mutants displaying that fitness level on the *y* axis. The metabolic relevance of a replicon in a particular environment is represented by the number of mutants showing phenotypes between the two extremes (1 and 0); the greater the metabolic relevance, the greater the number of non-extreme phenotypes.

Discussion

We have completed a comprehensive, manually and experimentally curated genome-scale metabolic reconstruction of a model multipartite genome of the N_2-fixing endosymbiont *S. meliloti*, and modelled the metabolic changes associated with niche transition. The switch from bulk soil to the rhizosphere was accompanied by a metabolic fine-tuning, primarily through changes in carbon metabolism and amino-acid biosynthesis. In contrast, moving from the rhizosphere to the nodule involved a comprehensive metabolic reprogramming. This involved essentially shutting off production of all biomass compounds and instead synthesizing co-factors necessary for a successful symbiosis, maximizing ATP production and fixing atmospheric nitrogen.

The analysis of the *in silico* fitness contributions of genes included in iGD1575 revealed that the chromosome is not metabolically specialized for a particular niche, but instead encodes the core metabolic machinery that enables growth of *S. meliloti* as a free-living microbe. In contrast, the evidence indicated that pSymB is metabolically specialized for the rhizosphere, helping *S. meliloti* to adapt to this environment and utilize the newly available substrates. The analysis failed to detect any environment where pSymA contributed to improved fitness, but it was seen that pSymA functions were solely relevant to the symbiotic process. Concerning multipartite genome evolution, these observations are consistent with an evolutionary scenario where (1) the gain of pSymB first significantly improved the ability of *S. meliloti* to colonize the rhizosphere as suggested previously[4,55], (2) pSymB gained additional genes encoding metabolic functions that contribute to fitness predominately in the rhizosphere and (3) pSymA only contributes metabolic functions relevant for establishing a N_2-fixing symbiosis.

We speculate that our observations here with *S. meliloti* may be generalizable to other bacteria with a multipartite genome that interact with a eukaryotic host. We hypothesize that secondary replicons might facilitate the start of a host interaction; this is the case for the large *Escherichia coli* virulence plasmids[7] and the rhizobial symbiotic plasmids[18]. Once the organism begins inhabiting the host-associated niche, the secondary replicon might acquire genes that improve fitness specifically in this new environment, whereas the chromosome remains largely undifferentiated, carrying the general metabolic pathways required for life and traits specific to the cell's original environment.

The modelling framework we have developed for this work can be adapted to study other types of biological association (for example, pathogenesis) and the metabolic reprogramming that is needed to operate the switch towards a novel ecological niche. Moreover, by demonstrating here that chromids and megaplasmids carry genes that primarily improve fitness in a specific niche, such as host interaction, this work illustrates secondary replicons as a rich reservoir of genes that have potential in synthetic biology applications. Finally, we anticipate that the iGD1575 model herein reconstructed will represent a valuable platform for future manipulations of *S. meliloti* aimed at its biotechnological exploitation in the context of agricultural procedures.

Methods

Metabolic network reconstruction. A draft metabolic model was constructed using the KBase Narrative Interface (www.kbase.us) and then manually and experimentally validated and expanded based on published data as described in the Supplementary Methods. The final *S. meliloti* model was termed iGD1575 in accordance with the nomenclature standard[56], and includes 1,575 genes, 1,825 reactions and 1,579 metabolites. The SBML file of the model was validated by the online SBML validator tool (http://sbml.org/Facilities/Validator/), and is available as Supplementary Data 6. Metabolic modelling was performed using Matlab R2015a (Mathworks), using scripts from the Cobra Toolbox[57] and the Gurobi 6.0.1 solver (www.gurobi.com). A detailed description of the modelling procedure is reported in the Supplementary Methods. For comparison of iGD1575 with previously published flux data[23], the flux through each reaction was constrained by setting the upper and lower bounds to the average plus or minus the error of the experimentally derived values. To facilitate construction of the *in silico* large-scale *S. meliloti* gene deletion mutants in iGD1575, identification of essential model genes was performed as described in Supplementary Note 6 and the essential iGD1575 genes are listed in Supplementary Table 7.

Biomass composition. No comprehensive description of the macromolecular composition of the *S. meliloti* biomass exists in the literature. However, such data are available for *Rhodobacter sphaeroides*, a related α-proteobacterium[58]. We therefore approximated the *S. meliloti* gross biomass composition using that of *R. sphaeroides*. Nevertheless, the specific composition of DNA, RNA, protein and lipids was adjusted based on the *S. meliloti* GC content[13], codon usage[59] and lipid composition[60–63]. Furthermore, succinoglycan was included in the biomass at 5% of the dry weight, which was estimated based on the literature[64–66]. The complete biomass composition is given in Supplementary Table 2.

Objective function formulation. The objective function for growth in synthetic media and bulk soil was set as a biomass reaction, producing biomass as described in the above section and fully detailed in Supplementary Table 2; this objective function was termed 'biomass_bulk_c0'. The objective function for growth in the rhizosphere (biomass_rhizo_c0) was the same as for bulk soil except that the amount of succinoglycan was doubled to account for biofilm formation on the plant root, and Nod factor was included (1 mg per g dry weight) as its production would be stimulated by the legume and is required for the initiation of symbiosis (Supplementary Table 2). Finally, the 'symbiosis_c0' objective function was adapted from a published *S. meliloti* model[19], and was used for modelling symbiosis. In short, the symbiosis objective function involved the synthesis of biomolecules relevant to symbiosis, as well as the export of L-alanine, L-aspartate and ammonium from fixed N_2.

***In silico* environmental representations.** *In silico* representations of the nutritional composition of the rhizosphere and bulk soil were constructed from data available in the literature (Table 3). For both soil representations, ammonium and nitrate were included at a one to one ratio, and sufficient ammonium, nitrate, phosphate, sulphate, metal ions and gases were included so that these compounds were not growth rate limiting. The relative abundance of the major carbon compounds was derived from the available literature as described in the Supplementary Methods. The boundaries of the exchange reactions used to define each environment are listed in Supplementary Table 6, as are the flux rate through all active exchange reactions.

Gene functional analysis. The WebMGA webserver[67] was used to provide functional COG annotations (P value cutoff of 0.001) to each gene in the model. Between-replicon biases were determined after standardizing by the number of genes from each replicon in iGD1575. To perform the COG analyses of the genes associated with variable reactions during the transition between niches, the COG annotation for each gene associated with the variable reaction classes was extracted from the WebMGA output of the previous COG analysis. Biases were determined after standardizing by the number of genes in each class of variable genes. Statistical significance was determined using Pearson's χ^2-tests. The complete list of COG annotations is available as Supplementary Data 7.

Phenotype MicroArray analysis. Phenotype MicroArray experiments using Biolog plates PM1 and PM2A were performed largely as described previously[25,49] with details elaborated on in the Supplementary Methods. All bacterial strains used in this study were described previously[31,49] and are listed in Supplementary Table 8. Of note, whereas most strains were inoculated from agar plates, *S. meliloti* RmP2754 (ΔB180) and a second wild-type control were inoculated from liquid M9-glucose cultures as RmP2754 grew poorly when inoculated directly from an agar plate. Data analysis was performed with DuctApe[27]. Activity index (AV) values were calculated following subtraction of the blank well from the experimental wells, whereas plots of the growth curves are of the unblanked data. Growth with each compound was considered positive if the AV value was ≥ 4. Negative growth phenotypes of the mutant strains were called if the AV value was ≤ 3, and only following manual inspection of the unblanked curves. However, it must be noted that a growth cutoff of 4 is likely to falsely eliminate some compounds that support slow growth of *S. meliloti*[27], such as beta-hydroxybutyrate (AV value = 3) and acetoacetate (AV value = 2)[68].

Growth curves and phosphate determination. *S. meliloti* was grown overnight in LBmc complex medium[4]. These cultures were washed with 0.85% saline and resuspended to an OD_{600} of ~0.05 in MM9 minimal medium[4] with either 7.5 mM glucose or 20 mM succinate as the sole carbon source. A volume of 200 µl of the cell suspensions were transferred in triplicate to wells of a 96-well microtitre plates and grown for 24 h at 30 °C with shaking in a BioTek Cytation 3 plate reader. OD_{600} readings were measured every 15 min. Growth rates were calculated between OD_{600} readings (not corrected for pathlength) of 0.1–0.5 with a previously developed Perl script[4].

To measure the amount of phosphate used by *S. meliloti*, the phosphate concentrations in the spent media following the completion of the growth curves were determined via the molybdenate blue—ascorbic acid colorimetric method[69]. In brief, cultures were centrifuged, 50 µl of supernatant was diluted with 5 ml of phosphate-free water and 0.8 ml of mixed reagent[69] was added to each sample. Following 10–30 min of incubation at room temperature, the A_{880} of each sample was measured and compared with a standard curve. The amount of phosphate remaining in spent media was compared with the phosphate present in the bacteria-free cultures to determine the amount of phosphate used by the bacteria. As the carbon source is growth limiting in these media, the used phosphate to carbon source ratio was calculated by dividing the amount of phosphate removed from the medium by the initial concentration of the carbon source (that is, 7.5 mM glucose or 20 mM succinate).

Data availability. The authors declare that the data supporting the findings of this study are available within the article and its Supplementary Information files. Matlab scripts used for generation of the FBA data are available from the authors on request.

References

1. Harrison, P. W., Lower, R. P. J., Kim, N. K. D. & Young, J. P. W. Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol.* **18,** 141–148 (2010).
2. Landeta, C. *et al.* Plasmids with a chromosome-like role in rhizobia. *J. Bacteriol.* **193,** 1317–1326 (2011).
3. Galardini, M., Pini, F., Bazzicalupo, M., Biondi, E. G. & Mengoni, A. Replicon-dependent bacterial genome evolution: the case of *Sinorhizobium meliloti*. *Genome Biol. Evol.* **5,** 542–558 (2013).
4. diCenzo, G. C., MacLean, A. M., Milunovic, B., Golding, G. B. & Finan, T. M. Examination of prokaryotic multipartite genome evolution through experimental genome reduction. *PLoS Genet.* **10,** e1004742 (2014).
5. Rocha, E. P. C. The organization of the bacterial genome. *Annu. Rev. Genet.* **42,** 211–233 (2008).
6. diCenzo, G. C., Zamani, M., Milunovic, B. & Finan, T. M. Genomic resources for identification of the minimal N2-fixing symbiotic genome. *Environ. Microbiol.* doi:10.1111/1462-2920.13221 (2016).
7. Johnson, T. J. & Nolan, L. K. Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* **73,** 750–774 (2009).
8. Morton, E. R., Merritt, P. M., Bever, J. D. & Fuqua, C. Large deletions in the pAtC58 megaplasmid of *Agrobacterium tumefaciens* can confer reduced carriage cost and increased expression of virulence genes. *Genome Biol. Evol.* **5,** 1353–1364 (2013).
9. Romanchuk, A. *et al.* Bigger is not always better: transmission and fitness burden of ~1MB *Pseudomonas syringae* megaplasmid pMPPla107. *Plasmid* **73,** 16–25 (2014).
10. Xu, Q., Dziejman, M. & Mekalanos, J. J. Determination of the transcriptome of *Vibrio cholerae* during intraintestinal growth and midexponential phase *in vitro*. *Proc. Natl Acad. Sci. USA* **100,** 1286–1291 (2003).
11. Becker, A. *et al.* Global changes in gene expression in *Sinorhizobium meliloti* 1021 under microoxic and symbiotic conditions. *Mol. Plant Microbe Interact.* **17,** 292–303 (2004).
12. Galardini, M. *et al.* Evolution of intra-specific regulatory networks in a multipartite bacterial genome. *PLoS Comput. Biol.* **11,** e1004478 (2015).
13. Galibert, F. *et al.* The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293,** 668–672 (2001).
14. Galardini, M. *et al.* Exploring the symbiotic pangenome of the nitrogen-fixing bacterium *Sinorhizobium meliloti*. *BMC Genomics* **12,** 235 (2011).
15. Hinsinger, P., Bengough, A. G., Vetterlein, D. & Young, I. M. Rhizosphere: biophysics, biogeochemistry and ecological relevance. *Plant Soil* **321,** 117–152 (2009).
16. Oldroyd, G. E. & Dixon, R. Biotechnological solutions to the nitrogen problem. *Curr. Opin. Biotechnol.* **26,** 19–24 (2014).
17. Geddes, B. A. *et al.* Use of plant colonizing bacteria as chassis for transfer of N_2-fixation to cereals. *Curr. Opin. Biotechnol.* **32,** 216–222 (2015).
18. Remigi, P., Zhu, J., Young, J. P. W. & Masson-Boivin, C. Symbiosis within symbiosis: evolving nitrogen-fixing legume symbionts. *Trends Microbiol.* **24,** 63–75 (2016).
19. Zhao, H., Li, M., Fang, K., Chen, W. & Wang, J. *In silico* insights into the symbiotic nitrogen fixation in *Sinorhizobium meliloti* via metabolic reconstruction. *PLoS ONE* **7,** e31287 (2012).
20. Schellenberger, J., Park, J. O., Conrad, T. M. & Palsson, B. Ø. BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions. *BMC Bioinformatics* **11,** 213 (2010).
21. Resendis-Antonio, O., Reed, J. L., Encarnación, S., Collado-Vides, J. & Palsson, B. Ø. Metabolic reconstruction and modeling of nitrogen fixation in *Rhizobium etli*. *PLoS Comput. Biol.* **3,** 1887–1895 (2007).
22. Resendis-Antonio, O. *et al.* Systems biology of bacterial nitrogen fixation: high-throughput technology and its integrative description with constraint-based modeling. *BMC Syst. Biol.* **5,** 120 (2011).
23. Fuhrer, T., Fischer, E. & Sauer, U. Experimental identification and quantification of glucose metabolism in seven bacterial species. *J. Bacteriol.* **187,** 1581–1590 (2005).
24. Yurgel, S., Mortimer, M. W., Rogers, K. N. & Kahn, M. L. New substrates for the dicarboxylate transport system of *Sinorhizobium meliloti*. *J. Bacteriol.* **182,** 4216–4221 (2000).
25. Biondi, E. G. *et al.* Metabolic capacity of *Sinorhizobium* (*Ensifer*) *meliloti* strains as determined by Phenotype MicroArray analysis. *Appl. Environ. Microbiol.* **75,** 5396–5404 (2009).
26. Spini, G. *et al.* Effect of the plant flavonoid luteolin on *Ensifer meliloti* 3001 phenotypic responses. *Plant Soil* **399,** 159–178 (2015).
27. Galardini, M. *et al.* DuctApe: a suite for the analysis and correlation of genomic and OmniLog Phenotype Microarray data. *Genomics* **103,** 1–10 (2014).
28. Fondi, M. *et al.* Genome-scale metabolic reconstruction and constraint-based modelling of the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125. *Environ. Microbiol.* **17,** 751–766 (2014).
29. Bartell, J. A., Yen, P., Varga, J. J., Goldberg, J. B. & Papin, J. A. Comparative metabolic systems analysis of pathogenic *Burkholderia*. *J. Bacteriol.* **196,** 210–226 (2014).
30. Schatschneider, S. *et al.* Establishment, *in silico* analysis, and experimental verification of a large-scale metabolic network of the xanthan producing *Xanthomonas campestris* pv. campestris strain B100. *J. Biotechnol.* **167,** 123–134 (2013).
31. Milunovic, B., diCenzo, G. C., Morton, R. A. & Finan, T. M. Cell growth inhibition upon deletion of four toxin-antitoxin loci from the megaplasmids of *Sinorhizobium meliloti*. *J. Bacteriol.* **196,** 811–824 (2014).
32. Cheng, J., Poduska, B., Morton, R. A. & Finan, T. M. An ABC-type cobalt transport system is essential for growth of *Sinorhizobium meliloti* at trace metal concentrations. *J. Bacteriol.* **193,** 4405–4416 (2011).
33. Malhotra, H. C., Prakash, J. & Sharma, G. C. Kinetics of chelation of Co(II) with L-histidine. *Proc. Indian Natl Sci. Acad.* **53,** 223–231 (1986).
34. Lerivrey, J. *et al.* Formation of D-glucosamine complexes with Cu(II), Ni(II) and Co(II) ions. *Inorg. Chim. Acta* **125,** 187–190 (1986).
35. Mauchline, T. H. *et al.* Mapping the *Sinorhizobium meliloti* 1021 solute-binding protein-dependent transportome. *Proc. Natl Acad. Sci. USA* **103,** 17933–17938 (2006).
36. Yamada, T., Letunic, I., Okuda, S., Kanehisa, M. & Bork, P. iPath2.0: interactive pathway explorer. *Nucleic Acids Res.* **39,** W412–W415 (2011).
37. Schmid, J., Sieber, V. & Rehm, B. Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. *Front. Microbiol.* **6,** 496 (2015).
38. Carlson, R. W., Price, N. P. & Stacey, G. The biosynthesis of rhizobial lipo-oligosaccharide nodulation signal molecules. *Mol. Plant Microbe Interact.* **7,** 684–695 (1994).
39. Capela, D., Filipe, C., Bobik, C., Batut, J. & Bruand, C. *Sinorhizobium meliloti* differentiation during symbiosis with alfalfa: a transcriptomic dissection. *Mol. Plant Microbe Interact.* **19,** 363–372 (2006).
40. Barnett, M. J., Toman, C. J., Fisher, R. F. & Long, S. R. A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote–host interaction. *Proc. Natl Acad. Sci. USA* **101,** 16636–16641 (2004).
41. Djordjevic, M. A. *Sinorhizobium meliloti* metabolism in the root nodule: a proteomic perspective. *Proteomics* **4,** 1859–1872 (2004).
42. Roux, B. *et al.* An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *Plant J.* **77,** 817–837 (2014).
43. Prell, J., Boesten, B., Poole, P. & Priefer, U. B. The *Rhizobium leguminosarum* bv. *viciae* VF39 γ-aminobutyrate (GABA) aminotransferase gene (*gabT*) is induced by GABA and highly expressed in bacteroids. *Microbiology* **148,** 615–623 (2002).
44. Geddes, B. A. & Oresnik, I. J. Physiology, genetics, and biochemistry of carbon metabolism in the alphaproteobacterium *Sinorhizobium meliloti*. *Can. J. Microbiol.* **60,** 491–507 (2014).
45. Udvardi, M. & Poole, P. S. Transport and metabolism in legume-rhizobia symbioses. *Annu. Rev. Plant Biol.* **64,** 781–805 (2013).
46. diCenzo, G. C. & Finan, T. M. Genetic redundancy is prevalent within the 6.7 Mb *Sinorhizobium meliloti* genome. *Mol. Genet. Genomics* **290,** 1345–1356 (2015).
47. diCenzo, G. C., Zamani, M., Cowie, A. & Finan, T. M. Proline auxotrophy in *Sinorhizobium meliloti* results in a plant-specific symbiotic phenotype. *Microbiology* **161,** 2341–2351 (2015).
48. Poysti, N. J., Loewen, E. D. M., Wang, Z. & Oresnik, I. J. *Sinorhizobium meliloti* pSymB carries genes necessary for arabinose transport and catabolism. *Microbiology* **153,** 727–736 (2007).
49. diCenzo, G., Milunovic, B., Cheng, J. & Finan, T. M. The tRNAarg gene and *engA* are essential genes on the 1.7-mb pSymB megaplasmid of *Sinorhizobium meliloti* and were translocated together from the chromosome in an ancestral strain. *J. Bacteriol.* **195,** 202–212 (2013).

50. Ramachandran, V. K., East, A. K., Karunakaran, R., Downie, J. A. & Poole, P. S. Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics. *Genome Biol.* **12,** R106 (2011).
51. Garcia-Fraile, P. *et al.* Arabinose and protocatechuate catabolism genes are important for growth of *Rhizobium leguminosarum* biovar *viciae* in the pea rhizosphere. *Plant Soil* **390,** 251–264 (2015).
52. MacLean, A. M., MacPherson, G., Aneja, P. & Finan, T. M. Characterization of the β-ketoadipate pathway in *Sinorhizobium meliloti*. *Appl. Environ. Microbiol.* **72,** 5403–5413 (2006).
53. MacLean, A. M., Haerty, W., Golding, G. B. & Finan, T. M. The LysR-type PcaQ protein regulates expression of a protocatechuate-inducible ABC-type transport system in *Sinorhizobium meliloti*. *Microbiology* **157,** 2522–2533 (2011).
54. Yurgel, S. N., Mortimer, M. W., Rice, J. T., Humann, J. L. & Kahn, M. L. Directed construction and analysis of a *Sinorhizobium meliloti* pSymA deletion mutant library. *Appl. Environ. Microbiol.* **79,** 2081–2087 (2013).
55. Egan, E. S., Fogel, M. A. & Waldor, M. K. Divided genomes: negotiating the cell cycle in prokaryotes with multiple chromosomes. *Mol. Microbiol.* **56,** 1129–1138 (2005).
56. Reed, J. L., Vo, T. D., Schilling, C. H. & Palsson, B. Ø. An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol.* **4,** R54 (2003).
57. Schellenberger, J. *et al.* Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.* **6,** 1290–1307 (2011).
58. Imam, S. *et al.* iRsp1095: a genome-scale reconstruction of the *Rhodobacter sphaeroides* metabolic network. *BMC Syst. Biol.* **5,** 116 (2011).
59. Nakamura, Y., Gojobori, T. & Ikemura, T. Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* **28,** 292 (1999).
60. Weissenmayer, B., Gao, J. L., López-Lara, I. M. & Geiger, O. Identification of a gene required for the biosynthesis of ornithine-derived lipids. *Mol. Microbiol.* **45,** 721–733 (2002).
61. Gao, J.-L. *et al.* Identification of a gene required for the formation of lyso-ornithine lipid, an intermediate in the biosynthesis of ornithine-containing lipids. *Mol. Microbiol.* **53,** 1757–1770 (2004).
62. Zavaleta-Pastor, M. *et al. Sinorhizobium meliloti* phospholipase C required for lipid remodeling during phosphorus limitation. *Proc. Natl Acad. Sci. USA* **107,** 302–307 (2010).
63. Basconcillo, L. S., Zaheer, R., Finan, T. M. & McCarry, B. E. A shotgun lipidomics study of a putative lysophosphatidic acid acyl transferase (PlsC) in *Sinorhizobium meliloti*. *J. Chromatogr. B* **877,** 2873–2882 (2009).
64. Wang, C. *et al.* Roles of poly-3-hydroxybutyrate (PHB) and glycogen in symbiosis of *Sinorhizobium meliloti* with *Medicago* sp. *Microbiology* **153,** 388–398 (2007).
65. Dorken, G., Ferguson, G. P., French, C. E. & Poon, W. C. K. Aggregation by depletion attraction in cultures of bacteria producing exopolysaccharide. *J. R. Soc. Interface* **9,** 3490–3502 (2012).
66. Glenn, S. A., Gurich, N., Feeney, M. A. & González, J. E. The ExpR/Sin quorum-sensing system controls succinoglycan production in *Sinorhizobium meliloti*. *J. Bacteriol.* **189,** 7077–7088 (2007).
67. Wu, S., Zhu, Z., Fu, L., Niu, B. & Li, W. WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics* **12,** 444 (2011).
68. Charles, T. C., Cai, G. Q. & Aneja, P. Megaplasmid and chromosomal loci for the PHB degradation pathway in *Rhizobium (Sinorhizobium) meliloti*. *Genetics* **146,** 1211–1220 (1997).
69. Murphy, J. & Riley, J. P. A modified single solution method for the determination of phosphate in natural waters. *Anal. Chim. Acta* **27,** 31–36 (1962).

Acknowledgements

GCD was supported by the Natural Science and Engineering Research Council of Canada (NSERC) through CGS-D and CGS-MSFSS awards. Work in the TMF lab is supported by NSERC.

Author contributions

G.C.D., M.F., A.M., M.G. and T.M.F. conceived the study; G.C.D. drafted, curated and validated the metabolic model; G.C.D. and M.F. performed the analyses on the model; L.D., M.G. and A.C. participated in model validation; G.C.D., M.F. and A.M. prepared the first draft of the manuscript; M.B. and C.V. contributed to drafting the manuscript. All authors read and approved the final version of the manuscript

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: diCenzo, G. C. *et al.* Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*. *Nat. Commun.* 7:12219 doi: 10.1038/ncomms12219 (2016).

Chapter 5

5.1 Conclusions

By now, it is known that many bacteria are able to establish a large variety of associations with plants, colonizing their surface, their tissues and/or forming with them symbiotic relationships. Among bacteria engaged for long term interactions with plants, those able to form nitrogen fixing symbiosis with legumes are collectively called rhizobia. As endophytic association, it is possible that more than one bacterial species could colonize the plant and form symbiotic root nodules, explaining the persistence of different rhizobial phenotypes during the evolution.

The most used legume-rhizobia symbiosis study model is the association established between *Sinorhizobium meliloti* and *Medicago*. During years, this symbiosis was explored by different points of view. The molecular and biochemical features which characterized the symbiotic association were largely explored and deepen. Then, lately, the exploration focus in particular on the different type of sociomicrobiological association established by bacteria with other bacteria or the plant. The topics of mixed nodules (more than one strain inside the same nodule) inspires the interest of scientists, but the real evidence of their existence it was not clearly demonstrated.

Therefore, in the second chapter of this thesis, we gave proof of the existence of mixed infections in nodules in the *Sinorhizobium meliloti- Medicago sativa* symbiosis, and the possibility of them to contain non-fixing strains. This demonstration can explain the

persistence of different symbiotic phenotypes during evolution, both mutualist and non-mutualists, despite the sanctions imposed by plant to the ineffective nodules. Indeed, it seems that the non-mutualist strains could be protected by plant sanctions masking by the mutualist strains, which provide a sufficient amount of fixed nitrogen to the legume. Neverthless, the impact and the role of the cheaters (non-fixing strains) on rhizobial evolution is hard to evaluate, in particular for the understanding of the fitness advantage they give to the plant.

In the last decades, also the interest and the exploration on the genomic evolutionary mechanisms behind the symbiotic association increased a lot, considering the peculiar multipartite genome composition of *Sinorhizobium* species. The function of a large number of dispensable genes are indeed not totally clarified, in particular in relation with their genome location, considering the hypothesis of a specific environmental function of each replicon.

In the third chapter of this thesis, we focus on the functional and phylogenetic profile of the gene *acdS* in *S. meliloti* strains, which is located on pSymA. The gene encodes the enzyme ACC deaminase, which is postulated to be involved degradation of the ethylene direct precursor, and it is present in many PGP bacteria.

The genomic context analysis in which the gene is included, showed the presence of mobile genetic elements in its close proximity among *S. meliloti* strains and in some other rhizobia. Accordingly, the phylogenetic analysis showed two different clades for *S. meliloti* strains, confirming a possible polyphyletic origin of *acdS*. Our results didn't evidence an involvement of the gene in the reduction of plant ethylene production and in the increasing of the symbiotic

competitiveness of *acdS* expressing strains. However, the gene seems to confer the ability to use unusual nitrogen sources, leaving to hypothesize a link with the increase of fitness in non-symbiotic conditions, (as endophytic lifestyle and rhizosphere colonization) which can be confirmed by the study of *acdS* expression profile.

Then, could the localization of *S meliloti acdS* be sintom of an initial evolutionary phase of genes acquisition by the megaplasmid?

Since the 80s, the multipartite organization of *S. meliloti* genome, and of a large number of bacteria, has aroused the interest of scientists. The three replicons which compounds *S. meliloti* genome, are characterized by an arrangement and a gene localization evolved over billion years, and therefore, their co-evolution implies strong metabolic connections among them. The actual evolutionary scenario, which describes a following acquisition of secondary replicons and of few housekeeping genes in the chromid, was enriched by the robustness of the metabolic pathways which connect the replions and by the hypothesis of secondary replicon's niche specialization.

In the fourth chapter, we perform *in silico* reconstruction of the *S. meliloti* cell metabolisms. Our results confirm the role of each replicon for the ecological niches inhabited by the bacterium, bulk soil, rhizosphere and symbiotic root nodule. In particular, the genome scale metabolic model predicts the differential activation of the metabolic pathways in the different environments, which are similar in the bulk soil and in the rhizosphere but very different in the nodule. Furthermore, the analysis showed that genes located in different replicons contribute differentially to the fitness of the bacterium in different conditions. From our results, accordingly to

the proposed evolutionary model, it seems that the chromosome is not metabolically specialized for particular niche. Instead, pSymB is metabolically specialized for the rhizosphere, helping *S. meliloti* to adapt to this environment and utilize the newly available substrates and pSymA functions are solely relevant to the symbiotic process.

In conclusions, in this thesis, three main aspects of the lifestyle of the facultative-symbiotic bacterium *S. meliloti* were broached. The sociomicrobiology topic, which usually comprises the study of the interactions between the rhizobia and the host legume and among different bacterial strains in the same plant, here is addressed toward the study of the cooperation established by bacterial strains with different symbiotic phenotype, which colonize the same root nodule. The molecular and functional analysis of the dispensable *S. meliloti* gene *acdS* is instead focused on the undestending of its phylogenetic history and its role, in particular in relation with the symbiotic interactions with the host plant. This study could be placed within a broader framework, aimed to a deeper comprehension of the functional role of accessory genes in relation with their location in the genome, accounting its multicomposite nature. Then, the third part of the thesis describes the *in silico* genome scale reconstruction of *S. meliloti* metabolism, to evaluate the contribute of each replicon to the microbial life cycle. This part is particularly important for the understanding of the evolutionary mechanism at the basis of the replicons specific niche-adaptation, and, to close the cycle, of the developed symbiotic lifestyle and cooperative phenotype.

In conclusion, the results presented in the four chapter of this thesis are aimed at providing a momentum for the understanding of plant-rhizobium interaction in the framework of a systems biology view of rhizobium symbiotic interaction, in which genome structure and

the plethora of dispensable genes in the pangenome may be used to predict the way of interaction among strains and plants (symbionts sociomicrobiology) and metabolic adaptation to the symbiosis. In perspective, the data *S. meliloti* reported here pose the bases for a predictive modelling of rhizobial symbiotic behavior in natural strains. Such modelling amy allowing to improve biotechnologically-related issues of symbiotic nitrogen fixation (e.g. élite strains development and selection) and allow to understand the ecology and evolution of plant-microbe interactions.

Appendix

6.1 Is the plant-associated microbiota of *Thymus* spp. adapted to the plant essential oil?

-Endophytes as Source of Bioactive molecules

In the last decades, the interest on the endophytic association between plant and its microbiota and their biological and chemical system, is increased a lot. This great attention is mainly based on the capability of endophytes to produce a great number of molecules with different biological activities. Furthermore, even the association plant-endophytes itself could represent a rich source of biologically active metabolites with an effect on the most disparate organisms. These natural products may have a great potential not only in agrochemical and biotechnology industries, but also in the pharmaceutical field (Christina et al., 2013). In this purpose, the increasing demand of novel source of antibiotics to contrast the bacterial multidrug resistance could find a response. Indeed, many studies of the last twenty years focused on the possible application of the endophytic products as antimicrobial compounds, but also as source of anticancer, antioxidant and immunosuppressive agents.

These applications are not only related with the molecules of bacterial production, but often the biochemical interaction endophytes-plant may assume a key role.

-Medicinal plants with essential oils and endophytes

As microbes, plants are able to produce a large variety of functional relevant metabolites, which, have been demonstrated to exhibit a diversity of medicinal properties. Most of these phytochemicals compounds are used by plant as defense against pathogens or competitors. Others, are instead represented by essential oils (EO) phenolic compounds, alkaloids, lectins/polypeptides and polyacetylenes (Cowan, 1999).

EO can be described a complex mixture of natural, volatile, and aromatic compounds synthesized by plants with medical use (Nerio et al., 2010). Many EOs biological activities have been reported such as antimicrobial, antiviral, antimycotic, antiparasitic, insecticidal, antidiabetic, antioxidant, and anticancer (Edris, 2007). The biological activity is mainly related with the chemical composition of the EO which is very variable and strongly dependent by the plant genotype and by the characteristics of its ecological niche. EOs can exhibit antibiotic potential against a large number of gram positive and gram negative bacteria. This is probably due to its extremely variable composition and to its lipophilic nature, which allows to destroy the cell wall and the cytoplasmatic membrane (Burt, 2004). By now, their commercial importance in the cosmetic, food, and pharmaceutical industries is well known. Despite the high relevance of medicinal plants producing antibacterial EOs, very little is known about their associated endophytic bacteria, even if some recent papers suggest their involvement in the production of plant bioactive molecules (Berg, 2005,Brader et al., 2014).

In this work, we investigate about the endophytic association between different species of the aromatic plants from the genus *Thymus* and their microbiota. In particular, we focus on the possible involvement of the microbiota in the EOs production, possibly related with their resistance profiles.

+ MODEL ARTICLE IN PRESS

ELSEVIER

Institut Pasteur

Research in Microbiology xx (2016) 1–7

www.elsevier.com/locate/resmic

Original Article

Is the plant-associated microbiota of *Thymus* spp. adapted to plant essential oil?

Alice Checcucci [a], Isabel Maida [a], Giovanni Bacci [a], Cristina Ninno [a], Anna Rita Bilia [b], Sauro Biffi [c], Fabio Firenzuoli [d], Guido Flamini [e], Renato Fani [a], Alessio Mengoni [a,*]

[a] *Department of Biology, University of Florence, Via Madonna del Piano 6, I-50019 Sesto Fiorentino, Florence 50019, Italy*
[b] *Department of Chemistry Ugo Schiff, University of Florence, Via Ugo Schiff 6, I-50019 Sesto Fiorentino, Florence 50019, Italy*
[c] *Il giardino delle Erbe, "Augusto Rinaldi Ceroni", Casola Valsenio, Ravenna, Italy*
[d] *Centre for Integrative Medicine, Careggi University Hospital, University of Florence, Florence, Italy*
[e] *Department of Pharmacy, University of Pisa, Via Bonanno Pisano 6, I-56126 Pisa, Italy*

Received 1 June 2016; accepted 11 November 2016
Available online ■ ■ ■

Abstract

We examined whether the microbiota of two related aromatic thyme species, *Thymus vulgaris* and *Thymus citriodorus,* differs in relation to the composition of the respective essential oil (EO). A total of 576 bacterial isolates were obtained from three districts (leaves, roots and rhizospheric soil). They were taxonomically characterized and inspected for tolerance to the EO from the two thyme species.

A district-related taxonomic pattern was found. In particular, high taxonomic diversity among the isolates from leaves was detected. Moreover, data obtained revealed a differential pattern of resistance of the isolates to EOs extracted from *T. vulgaris* and *T. citriodorus*, which was interpreted in terms of differing chemical composition of the EO of their respective host plants.

In conclusion, we suggest that bacterial colonization of leaves in *Thymus* spp. is influenced by the EO present in leaf glandular tissue as one of the selective forces shaping endophytic community composition.

Keywords: Plant microbiota; Medicinal plants; Essential oil; *Thymus*; Bacterial endophytes; Adaptation

1. Introduction

Plant-associated microbiota is constituted by a panoply of prokaryotic and/or eukaryotic microorganisms living in close contact with and within plant tissues [1]. The differences related to plant species and cultivar in the microbiota structure and composition were previously described [2]. Every plant species, as well as the different soil types and related rhizosphere, can be colonized by specific microbial communities. Moreover, the intimate association of microorganisms with plants has broad implications for plant nutrient assimilation, growth, stress tolerance and health status [3–5]. Depending on the colonized plant environment (internal tissues, leaves and root surface), endophytes, epiphytes and rhizospheric microorganisms are recognized, respectively. Generally, plant endophytes are thought to enter from roots (e.g. cracking at lateral root emergences) and/or leaves (e.g. from stomata) [1], or via insect bites (e.g. phytopathogens). However, cases of vertical transmission through seeds are also reported [6] which, in turn, imply a more intimate association and the possibility of co-evolution of plants with their microbiota.

A key question related to plant-associated bacteria is their possible adaptation to secondary metabolites produced by the plant, particularly to those exhibiting antibacterial activity. Indeed, plants accumulating toxic molecules, as, for instance hyperaccumulator plants [4,7], are enriched in bacteria resistant to these toxic compounds; however, very few data

* Corresponding author.
E-mail address: alessio.mengoni@unifi.it (A. Mengoni).

http://dx.doi.org/10.1016/j.resmic.2016.11.004

Please cite this article in press as: Checcucci A, et al., Is the plant-associated microbiota of *Thymus* spp. adapted to plant essential oil?, Research in Microbiology (2016), http://dx.doi.org/10.1016/j.resmic.2016.11.004

concerning this issue in other plant species have been reported up to now [8]. Plant species producing essential oils (EOs) could be good models for testing the hypothesis of a selection of plant microbiota in relation to the plant's secondary metabolites. In fact, EOs may have a strong antibacterial effect [9], and we hypothesize that plants producing antibacterial EOs may favor colonization of tissue(s) by bacteria resistant to these oils.

Among the most relevant plant species producing EOs with antibacterial activity, those belonging to the genus *Thymus* are of special interest [10–15] due to the presence of thymol, a monoterpene phenol isomeric, with carvacrol as one of the main possible active components. One of the most common *Thymus* species for EO extraction is *Thymus vulgaris* (or garden thyme), a wild plant quite common in Southern Europe. Different species of *Thymus* produce EOs with very different chemical compositions, especially in thymol content [11,12]. Among those species, *Thymus citriodorus* is one of the most interesting, since its EO has a very low thymol content and a high geraniol and citronellol content [16]. Consequently, *T. vulgaris* and *T. citriodorus*, though phylogenetically close, produce EOs having extremely different chemical profiles and might represent very good models for testing the hypothesis of a plant chemotype-specific microbiota.

Therefore, the aim of this work was to check possible relationships existing between EOs present in plant tissues and adaptation of bacterial microbiota. For this purpose, we characterized, at the taxonomic and phenotypic levels (in terms of resistance to EO, seen as an antimicrobial compound), a pool of 576 isolates randomly collected from leaves, roots and rhizospheric soil of *T. vulgaris* and *T. citriodorus*. This cultured microbiota was taxonomically characterized by 16S rRNA metagenomic analysis and tolerance to the EO produced by the respective host plant.

2. Materials and methods

2.1. Isolation and phenotypic characterization of T. vulgaris and T. citriodorus isolates

Three *T. vulgaris* and three *T. citriodorus* plants cultivated in soil in an open air common garden at the "Giardino delle Erbe", Casola Valsenio, Ravenna, Italy (44.23° N, 11.62° E), were collected in May 2014. For each of the six plants, 10 g of the rhizospheric soil (RS) and 10 g of the two plant anatomic parts (roots, R; leaves, L) were sampled. Hereafter, these three parts will be referred to as "districts". Duplicate samples were taken for each plant for a total of 36 samples (2 × 3 districts × 6 plants). Isolation of bacteria from each sample was carried out as reported in Chiellini et al. [17], including surface sterilization and washing of plant material to remove most epiphytic bacteria and their DNA. Furthermore, in this work, the samples were homogenized in a sterile mortar with the addition of a large volume (20 ml) of 0.9% NaCl (Sigma Aldrich, USA), in order to limit the possible antimicrobial effect of EOs released from crushed tissues. Following the previously reported protocol [17], fast-growing endophytic and rhizospheric bacteria were isolated by plating on solid tryptone soya broth (TSB) medium (BioRad, CA, USA) and incubating at 30 °C for 48 h. Isolation of EO-resistant bacteria was performed on solid TSB containing 0.025% v/v of EO from commercial white thyme (*T. vulgaris*) (Erboristeria Spada, Florence, Italy), *T. vulgaris* and *T. citriodorus.* This concentration of EO was chosen after preliminary analyses, indicating a medium selective effect (lower than the minimal inhibitory concentration, or MIC) toward a panel of bacterial strains (unpublished data) and a literature survey of antimicrobial activity against bacteria [18–20]. Bacterial load was expressed as colony-forming-units (CFUs)/g of fresh plant material or rhizospheric soil. Each CFU estimation was performed in duplicate for each sample. From all the isolation plates for each of the three districts (RS, R, L, for both plant species), 96 colonies were randomly picked, re-isolated on solid TSB medium and stored under TSB/glycerol (25% final glycerol concentration) at −80 °C. Consequently, two biobanks, each including a total of 576 isolates were constituted, representing the microbiota grown on solid TSB and on solid TSB containing 0.025% white thyme EO, respectively. Then, from each biobank, three pools representing the cultured microbiota from the three districts (R, RS, L) were prepared by mixing 10 µl of $OD_{600} = 1.5$ individual liquid cultures for each of the 96 isolates. A preliminary analysis on 20 randomly collected isolates indicated that $OD_{600} = 1.5$ corresponded to approximately $1 \cdot 10^9$ cells/ml for each isolate. These pools of the cultured microbiota were used for total DNA extraction. EO resistance tests were performed by plating the individual isolates of the two biobanks on solid TSB plates containing 0.025% v/v EOs extracted from the very same plants (*T. vulgaris* and *T. citriodorus*) and incubating at 30 °C. Growth was inspected after 48 h. Resistant isolates were defined as those for which growth was observed.

2.2. EO extraction and determination of essential oil composition

The same plants used for sampling RS, R and L microbiota were then pooled together for each species (*T. vulgaris* and *T. citriodorus*), and steam distillation of EO was carried out at the Giardino delle Erbe "Augusto Rinaldi Ceroni" Casola Valsenio (RA) (Italy), following the method described in [21]. We used the same weight of fresh plants (approximately 1 kg of above ground tissues), and the same amount (1.6 ml) was distilled from the two species. EO composition was determined following the GC method described in [22].

2.3. 16S rRNA gene metagenomic sequencing and analysis

DNA was extracted from the pooled cultured microbiota samples by using the FastDNA™ SPIN Kit for soil (MP Biomedicals, Italy). DNA extracted from the pools was analyzed by ethidium-bromide-stained agarose gel (0.8% TAE w/v) electrophoresis and spectrophotometrically quantified

Please cite this article in press as: Checcucci A, et al., Is the plant-associated microbiota of *Thymus* spp. adapted to plant essential oil?, Research in Microbiology (2016), http://dx.doi.org/10.1016/j.resmic.2016.11.004

using the Infinite® M200 PRO NanoQuant (Tecan, Milan, Italy). The bacterial V3–V4 hypervariable region of 16S rRNA genes was amplified from each sample using specific primers (5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG-3′ and 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC-3′) as previously reported [23]. The 16S metagenomic libraries were prepared according to Illumina MiSeq guidelines and sequenced at the IGA Technology Services (http://www.igatechnology.com/) (Udine, Italy) using the Illumina MiSeq technology with pair-end sequencing [24]. Obtained paired end reads were 300 ± 2 bp in length. Library preparation and demultiplexing were performed following Illumina's standard pipeline. Raw sequences, generated as described above, were trimmed with StreamingTrim [25] and processed through an automated O2tab Pipeline for "Operational Taxonomic Unit" (OTU) clustering of microbiome data (https://github.com/GiBacci/o2tab) following a previously described workflow [26]. OTUs were clustered at 97% of the sequence identity threshold. From the OTU (cluster) produced above, a single representative sequence was selected and used for taxonomical identification by SINA classifier on the latest SILVA dataset available when we performed the analysis (SSURef Nr99). Statistical analyses were performed on an OTU table, as previously reported [26,27], using R scripts and the R package Vegan [28].

Sequences have been deposited in the NCBI database under Bioproject PRJNA304275.

3. Results and discussion

3.1. T. citriodorus and *T. vulgaris*-associated microbiota are strongly different at the leaf level

The cultivable bacterial load in thyme rhizospheric soil, roots and leaves ranged from about 10^1 to 10^6 CFU/g of soil/root or leaves (fresh weight) (Fig. 1); these data are in agreement with previous reports on a number of different plant species (see for instance [7,17,29]). In both *T. vulgaris* and *T. citriodorus*, bacterial loads were higher in soil than in leaves, a trend in agreement with similar findings reported for other plant species (see references above). Moreover, significant differences between the two plant species were observed for bacterial loads in both roots and leaves. Indeed, *T. citriodorus* exhibited higher bacterial endophytes titers than *T. vulgaris*, especially in leaves, with the *T. citriodorus*/*T. vulgaris* bacterial load ratio approaching 3 log. This suggests that leaves of *T. vulgaris* are a more hostile environment to bacterial colonization than those of *T. citriodorus*. Indeed, the EO of *T. vulgaris*, which is produced by glandular tissue in leaves, is known to have strong antibacterial activity [11]. Consequently, it can be expected that leaves might be more toxic for bacteria than roots and rhizospheric soil, in agreement with the finding that the main differences between *T. citriodorus* and *T. vulgaris* in terms of bacterial load are shown by leaves.

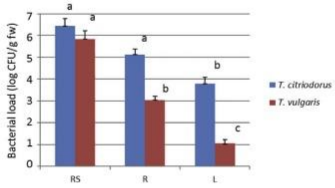


Fig. 1. Bacterial load of *T. citriodorus* and *T. vulgaris*. Bar chart reports mean values (± one standard deviation from the three biological replicates) of CFUs per gram of fresh weight. RS, rhizospheric soil; R, root; L, leaves. Different letters indicate significantly different values (one-way ANOVA, $P < 0.05$).

With the aim of determining whether such differential endophytic bacterial loads also run parallel to differences in the taxonomic composition of the pool of bacterial isolates (here referred to as "cultured microbiota") between plant districts and the two plant species, six pooled collections, each consisting of 96 randomly picked isolates (leaves, roots and rhizospheric soil for both *T. vulgaris* and *T. citriodorus*), were prepared for a total of 576 bacterial isolates. These pools of isolates were then taxonomically characterized by 16S rRNA gene metagenomic sequencing of the V3–V4 region. After trimming and quality filtering, 673,465 reads were used for OTU clustering and taxonomic attribution (Supplemental Table S1). As shown in Fig. S1, all rarefaction curves reached saturation, setting the OTU cut-off at 96% similarity, suggesting that bacterial biodiversity was adequately represented. The bioinformatic analysis performed as described in Materials and methods enabled us to recognize a total of 15 OTUs (Supplemental Table S2). After assigning OTUs to the bacterial taxonomy (Supplemental Table S2), 13 OTUs were attributed to a bacterial taxon, whereas 2 OTUs could not be assigned to any taxon, since they very likely were chimeric molecules. Three main phyla were recognized, *Proteobacteria* (7 OTUs), *Firmicutes* (5 OTUs) and *Actinobacteria* (1 OTU).

The presence of fast-growing taxa (e.g. *Bacillus*, *Pseudomonas*, Enterobacteriaceae) was very likely related to the selection used in this work for fast-growing copiotrophic bacteria; indeed, isolates were collected on TSB plates after two days of incubation at 30 °C. The dominance of Gammaproteobacteria in such culturable microbiota was in agreement with previous findings on culturable endophytic bacteria in several plant species [7,17,29–33]. A cluster analysis performed on the OTU table (Fig. 2) revealed that the microbiota from RS and R districts were more similar between them than with the microbiota isolated from the L district. The difference in leaves endophytic microbiota from RS and R was also obtained by analysis of the amplified 16S rRNA gene from environmental DNA (eDNA) extracted from the same plant tissues and rhizospheric soil used in the bacterial isolation procedure (See Supplementary Table S3, Supplementary Fig. S2).

These data suggested that the plate isolation procedure may have potentially influenced taxonomic differences among bacterial communities living in the different plant districts; in other words, it is possible that the composition of the isolated

Please cite this article in press as: Checcucci A, et al., Is the plant-associated microbiota of *Thymus* spp. adapted to plant essential oil?, Research in Microbiology (2016), http://dx.doi.org/10.1016/j.resmic.2016.11.004

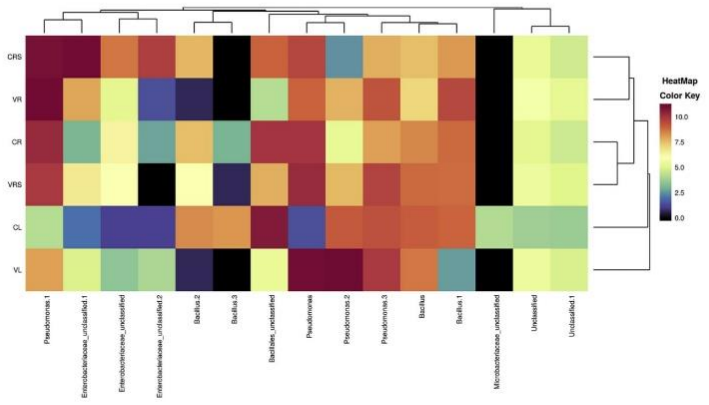


Fig. 2. Taxonomic composition and similarities of the microbiota. Heatmap of taxonomic assignments down to the genus level. Clustering of samples and taxonomic assignments after Bray–Curtis distance estimation are reported. C, *T. citriodorus*; V, *T. Vulgaris;* RS, rhizospheric soil; R, root; L, leaves.

culturable microbial communities might be related to the contents of thymol, geraniol, citronellol and/or other molecules present in the EO-rich leaves of *T. vulgaris* and *T. citriodorus* [11]. Indeed, although precautions were taken during the isolation procedure in order to limit the toxicity of EOs released from ground-up leaf tissues, it cannot be excluded that some of the less resistant endophytic bacteria living in the leaves were counter-selected by the release of EO from glandular tissue.

To evaluate taxa that mainly contribute to differentiating between *T. vulgaris* and *T. citriodorus* microbiota, a Simper analysis was carried out. Data obtained are shown in Table 1; the analysis revealed that, as expected on the basis of their relative abundance, OTUs assigned to *Pseudomonas* were responsible for most of the differences existing between *T. vulgaris* and *T. citriodorus,* and between leaves and roots. Moreover, rhizospheric soil and roots were differentiated by members of *Enterobacteriacea.*

Table 1
Differentiation between *Thymus* ssp. bacterial communities.

OTU	Taxonomic assignment	% Contribution
T. citriodorus* vs. *T. vulgaris		
OTU 7	*Pseudomonas*	27.36
OTU 6	*Pseudomonas*	20.71
OTU 4	*Pseudomonas*	20.35
Leaves vs. roots		
OTU 7	*Pseudomonas*	28.70
OTU 4	*Pseudomonas*	20.57
OTU 6	*Pseudomonas*	18.18
Soil vs. roots		
OTU 10	Enterobacteriaceae unclassified	44.22
OTU 12	Enterobacteriaceae unclassified	13.41
OTU 7	*Pseudomonas*	12.31

Results of Simper test in OTU occurrence between *T. citriodorus* and *T. vulgaris* and between cultured microbiota in the different districts (plant organs as leaves and roots and rhizospheric soil) is reported. Values indicate the percentage of variance contributing to each OTU.

3.2. Is the plant-associated microbiota differentially adapted to the EO of the host plant?

To ascertain whether EOs synthesized in glandular tissue of leaves could constitute an environmental factor shaping plant-associated bacterial community composition, the cultured microbiota pools from *T. vulgaris* and *T. citriodorus* were checked for their tolerance to EOs extracted from the very same plants (*T. vulgaris* and *T. citriodorus*) and to a commercial preparation of thyme EO (i.e. white thyme, *T. vulgaris*). The latter was first tested as a reference. Data obtained are shown in Fig. 3, whose analysis revealed an increase in the percentage of white thyme-resistant *T. vulgaris*-associated bacterial isolates from soil to leaves. This trend was not observed for the cultivable bacteria isolated from *T. citriodorus.* Moreover, the percentage of resistant endophytic isolates in leaves was lower in *T. citriodorus* than in *T. vulgaris* ($P < 0.05$ for soil and leaves). These results are in agreement with the differential composition of EO extracted from *T. vulgaris* and *T. citriodorus* plant species, with respect to the EO of white thyme (Table 2). In fact, EO of *T. vulgaris* has three main components, geraniol and its derivative geranyl acetate (representing about 52.9% of the total constituents) and thymol (about 10.9%), which is the typical phenol derivative of *Thymus* species. White thyme EO contains ca. 43% of thymol and ca. 48% of an aromatic monoterpene, p-cymene. *T. citriodorus* EO lacks thymol, whereas it is rich in citronellol (the hydrogenated derivative of geraniol) and geranyl acetate. It is known that thymol exerts strong antimicrobial activity due

Please cite this article in press as: Checcucci A, et al., Is the plant-associated microbiota of *Thymus* spp. adapted to plant essential oil?, Research in Microbiology (2016), http://dx.doi.org/10.1016/j.resmic.2016.11.004

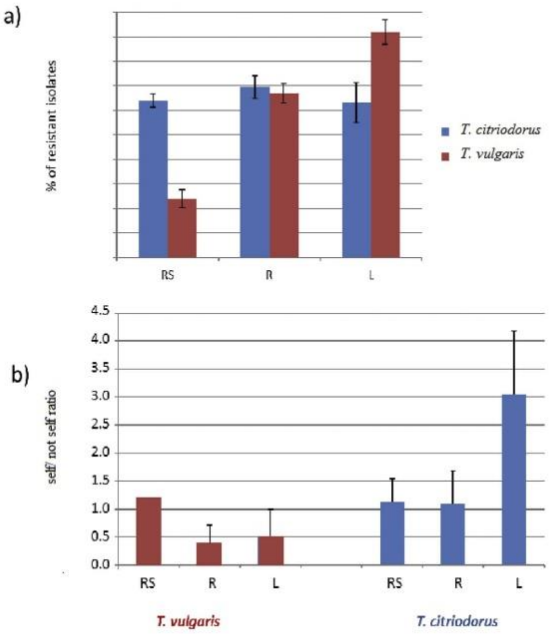


Fig. 3. Adaptation of microbiota to *thymus* EO. a) Percentage of resistance to commercial white thyme EO of cultured microbiota. RS, rhizospheric soil; R, root; L, leaves. b) Ratio between the percentage of microbiota resistant to self versus non-self EO in *T. vulgaris* (V) and *T. citriodorus* (C). RS, rhizospheric soil; R, root; L, leaves. Bar chart reports mean values (± one standard deviation) from triplicate experiments (three individual plants) for each sample. See text for details.

principally to the phenol moiety [34], which is not exhibited by citronellol and geranyl acetate [35]. Therefore, it might be expected that *T. vulgaris* leaf-associated bacteria, coping with the strong antimicrobial activity of *T. vulgaris* EO, could be more resistant to an EO based on similar chemical composition (white thyme) than *T. citriodorus*-associated endophytes.

Moreover, if EOs exert selective pressure on endophytic colonization, it can be suggested that the endophytic communities from a given plant species are more adapted (resistant) to the EO produced by the same plant species. To test this hypothesis, we analyzed the level of cross-resistance of *T. vulgaris*- and *T. citriodorus*-associated bacteria against the EO extracted from either the same or different plant species (*T. vulgaris* and *T. citriodorus*). Data obtained on the same panel of 576 isolates are shown in Fig. 3b. The ratio between the percentages of isolates resistant to the EO by the same plant species ('self') and the percentages of isolates resistant to the EO by the other plant species ('non-self') was measured. Hence, values >1 indicated that cultured microbiota is better adapted (i.e. resistant) toward the 'self' EO than toward the 'non-self' EO. On the other hand, values < 1 indicate greater resistance to 'non-self' EO than to 'self' EO.

Data shown in Fig. 3b reveal that values of *T. vulgaris* cultured microbiota were below 1 for both roots and leaves ($P < 0.05$); for *T. citriodorus,* strong tolerance toward 'self' EO was observed for leaf endophytes (ratio = 3.05 ± 1.13). These results can be interpreted in terms of adaptation of *T. citriodorus* leaf bacterial isolates to *T. citriodorus* EO. The tolerance for 'non-self' EO in *T. vulgaris* isolates can be interpreted in terms of the different chemical composition of the respective EOs. It is known that the main constituents of *T. vulgaris* EO (the phenols carvacrol and thymol) possess high antibacterial activity [11,36]. In contrast, *T. citriodorus* EO mainly contains aliphatic terpenoids. Further confirmation of this interpretation came from the same analysis performed with the fraction of isolates selected for resistance to white thyme EO (Supplemental Fig. S2). Here, *T.*

Please cite this article in press as: Checcucci A, et al., Is the plant-associated microbiota of *Thymus* spp. adapted to plant essential oil?, Research in Microbiology (2016), http://dx.doi.org/10.1016/j.resmic.2016.11.004

Table 2
Percentage of chemical compounds found in *T. vulgaris*, *T. citriodorus* and white thyme (commercial *T. vulgaris*) EOs.

Components	*T. vulgaris*	*T. citriodorus*	White thyme
1.4-Cineole	0.0	0.0	0.1
1.8-Cineole (=eucalyptol)	0.3	0.0	0.2
1-Octen-3-ol	0.5	0.5	0.0
3-Octanol	0.1	0.1	0.0
4-Terpineol	0.4	0.0	0.0
Alloaromadendrene	0.3	0.3	0.0
Bicyclogermacrene	2.8	1.2	0.0
Camphene	0.1	0.0	0.1
Camphor	0.8	0.9	0.0
Carvacrol	0.6	0.0	0.4
Caryophyllene oxide	0.8	1.1	Tr
Citronellol	0.0	**40.1**	0.0
Citronellyl acetate	0.1	0.2	0.0
Exo-fenchol	0.0	0.0	Tr
Geranial	4.4	5.7	0.0
Geraniol	**28.7**	0.0	0.0
Geranyl acetate	**24.2**	**27.5**	0.0
Geranyl formate	0.0	0.2	0.0
Geranyl isobutyrate	0.2	0.3	0.0
Geranyl n-butyrate	0.0	1.5	0.0
Geranyl propionate	2.5	2.9	0.0
Isoborneol	0.7	0.5	0.0
Isobornyl acetate	0.3	0.2	0.0
Limonene	0.0	0.0	0.2
Linalool	1.2	0.4	1.2
Methyl carvacrol	0.3	0.0	0.0
Myrcene	0.1	0.0	0.0
Neral	3.8	4.4	0.0
Nerol	2.3	3.9	0.0
Neryl propanoate	0.1	0.2	0.0
Neryl-acetate	0.9	1.4	0.0
p-Cymene	2.3	0.0	**47.9**
Spathulenol	0.5	1.3	0.0
Thymol	**10.9**	0.0	**43.1**
Thymol methyl ether	0.9	0.0	0.0
Trans-α-bergamotene	0.0	0.0	Tr
tricyclene	0.0	0.0	Tr
α-Humulene	0.1	0.0	Tr
α-Pinene	0.0	0.0	4.3
α-Terpinene	0.1	0.0	0.0
α-Terpineol	0.0	0.0	0.6
β-Bisabolene	1.5	1.9	0.0
β-Caryophyllene	3.9	2	0.2
β-Pinene	0.0	0.0	1.2
γ-Eudesmol	0.0	0.1	0.0
γ-Muurolene	0.3	0.0	0.0
γ-Terpinene	0.9	0.0	0.0
δ-3-Carene	0.0	0.0	Tr
δ-Cadinene	0.2	0.0	0.0

Abbreviation: Tr, traces.
In bold, the main constituents of each oil.

vulgaris leaf isolates display clear tolerance for 'self' EO, suggesting that, though *T. vulgaris* and white thyme EOs both contain thymol, *T. vulgaris* leaf endophytes better cope with EO from their host plant species.

3.3. Conclusions

Here we provide evidence that a link between a plant chemotype and its microbiota is present, in terms of the effect exhibited by leaf EOs on leaf microbiota. Indeed, our findings showed that the endophytic microbiota of two phylogenetically close species, *T. vulgaris* and *T. citriodorus*, producing EO with highly different chemical composition but both containing thymol, are strongly differentiated in terms of taxonomic composition and especially in relation to tolerance to the relative host plant EO. Previous work showed that plants exhibit different ecological niches in their organs, which might select for specific microbiota assemblages [7,31,32,37,38]. The finding that differences among microbiota and tolerance to the EOs were higher in leaf endophytic microbiota than in root microbiota suggests a tight correlation/co-occurrence of bacterial strains, with EO-producing organs. Hence, data obtained in this work support the hypothesis of selection of the plant microbiota in relation to the chemotype of the EO produced by the host plant, in particular for *T. vulgaris*. In other words, the composition of EO synthesized in the leaves of the plant itself may represent a key factor shaping the structure of cultivable bacterial communities living in different plant districts. According to this hypothesis: i) the cultivable microbiota isolated from leaves of the two plant species are less "similar" from a phylogenetically viewpoint than those isolated from roots and rhizospheric soil; ii) the microbiota from leaves of *Thymus citrodorus* is more sensitive to both *T. vulgaris* and white thyme EOs than the leaf microbiota of *T. vulgaris*; iii) the number of cultivable bacteria isolated from *T. citrodorus* leaves is about 3 orders of magnitude higher than that isolated from that from *T. vulgaris* leaves, suggesting the existence of selective pressure (i.e. EO activity) lower in *T. citrodorus* than in *T. vulgaris*. Otherwise, we cannot a priori completely exclude possible antimicrobial activity of the EO acting during sample preparation from fresh leaf tissue that, in turn, may have caused a bias in the OTU representation of cultured microbiota.

The involvement (both direct and indirect) of endophytic microorganisms in production of EO has been shown in several systems, including both bacteria and fungus-plant interactions [39]. Our findings lead us to speculate that adaptation of the bacterial communities to the EO could be related to involvement of bacteria in the metabolism of EO production by the plant. Therefore, the leaf glandular tissue and the production of EO can be considered as selective habitats and forces shaping endophytic (and possibly all plant-associated) bacterial microbiota.

Further investigation of tissue localization of selected endophytes in relation to glandular organs and their metabolic characterization should shed additional light on the possible interaction of leaf endophytic microbiota with EO production in *Thymus*.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

This work was supported by the Ente Cassa di Risparmio di Firenze (project 2013.0657).

Please cite this article in press as: Checcucci A, et al., Is the plant-associated microbiota of *Thymus* spp. adapted to plant essential oil?, Research in Microbiology (2016), http://dx.doi.org/10.1016/j.resmic.2016.11.004

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.resmic.2016.11.004.

References

[1] Reinhold-Hurek B, Hurek T. Living inside plants: bacterial endophytes. Curr Opin Plant Biol 2011;14:435–43.
[2] Berg G, Smalla K. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. FEMS Microbiol Ecol 2009;68(1):1–13.
[3] Vandenkoornhuyse P, Quaiser A, Duhamel M, Le Van A, Dufresne A. The importance of the microbiome of the plant holobiont. New Phytol 2015;206:1196–206.
[4] Mengoni A, Schat H, Vangronsveld J. Plants as extreme environments? Ni-resistant bacteria and Ni-hyperaccumulators of serpentine flora. Plant Soil 2010;331:5–16.
[5] Schlaeppi K, Bulgarelli D. The plant microbiome at work. Mol Plant Microbe Interact 2015;28:212–7.
[6] Truyens S, Weyens N, Cuypers A, Vangronsveld J. Bacterial seed endophytes: genera, vertical transmission and interaction with plants. Environ Microbiol Rep 2015;7:40–50.
[7] Barzanti R, Ozino F, Bazzicalupo M, Gabbrielli R, Galardi F, Gonnelli C, et al. Isolation and characterization of endophytic bacteria from the nickel hyperaccumulator plant *Alyssum bertolonii*. Microb Ecol 2007;53:306–16.
[8] Karamanoli K, Menkissoglu-Spiroudi U, Bosabalidis A, Vokou D, Constantinidou H. Bacterial colonization of the phyllosphere of nineteen plant species and antimicrobial activity of their leaf secondary metabolites against leaf associated bacteria. Chemoecology 2005;15(2):59–67.
[9] Edris AE. Pharmaceutical and therapeutic potentials of essential oils and their individual volatile constituents: a review. Phytotherapy Res 2007; 21(4):308–23.
[10] Pesavento G, Calonico C, Bilia AR, Barnabei M, Calesini F, Addona R, et al. Antibacterial activity of *Oregano*, *Rosmarinus* and *Thymus* essential oils against *Staphylococcus aureus* and *Listeria monocytogenes* in beef meatballs. Food Control 2015;54:188–99.
[11] Nabavi SM, Marchese A, Izadi M, Curti V, Daglia M, Nabavi SF. Plants belonging to the genus *Thymus* as antibacterial agents: from farm to pharmacy. Food Chem 2015;173:339–47.
[12] Ćavar Zeljković S, Maksimović M. Chemical composition and bioactivity of essential oil from *Thymus* species in Balkan Peninsula. Phytochem Rev 2015;14:335–52.
[13] Xue J, Michael Davidson P, Zhong Q. Antimicrobial activity of thyme oil co-nanoemulsified with sodium caseinate and lecithin. Int J Food Microbiol 2015;210:1–8.
[14] Mollarafie P, Khadiv Parsi P, Zarghami R, Amini Fazl M, Ghafarzadegan R. Antibacterial and wound healing properties of thymol (*Thymus vulgaris* Oil) and its application in a novel wound dressing. J Med Plants 2015;14:69–81.
[15] Magi G, Marini E, Facinelli B. Antimicrobial activity of essential oils and carvacrol and synergy of carvacrol and erythromycin, against clinical, erythromycin-resistant Group A Streptococci. Front Microbiol 2015:6.
[16] Omidbaigi R, Fattahi F, Alirezalu A. Essential oil content and constituents of *Thymus* × *citriodorus* L. at different phenological stages. J Essent Oil Bearing Plants 2009;12:333–7.
[17] Chiellini C, Maida I, Emiliani G, Mengoni A, Mocali S, Fabiani A, et al. Endophytic and rhizospheric bacterial communities isolated from the medicinal plants *Echinacea purpurea* and *Echinacea angustifolia*. Int Microbiol 2015;17:165–74.
[18] Sfeir J, Lefrançois C, Baudoux D, Derbré S, Licznar P. In vitro antibacterial activity of essential oils against *Streptococcus pyogenes*. Evid Based Complement Alternat Med 2013;2013:9. Article ID 269161.
[19] Sqalli H, El Ouarti A, Farah A, Ennabili A, Haggoud A, Ibnsouda S, Houari A, Houssaini Iraqui M. Antibacterial activity of *Thymus pallidus* Batt. and determination of the chemical composition of its essential oil. Acta Bot Gallica 2009;156(2):303–10.
[20] Souza Stamford, Lima Trajano, Brazilian BF. Antimicrobial effectiveness of spices: an approach for use in food conservation systems. Archives Biol Technol 2005;48(4):549–58.
[21] Hancı SS, Sahin Serpil, Yılmaz Levent. Isolation of volatile oil from thyme (*Thymbra spicata*) by steam distillation. Food/Nahrung 2003;47(4):252–5.
[22] Maida I, Lo Nostro A, Pesavento G, Barnabei M, Calonico C, Perrin E, et al. Exploring the anti- *Burkholderia cepacia* complex activity of essential oils: a preliminary analysis. Evid Based Complement Alternat Med 2014;2014: 10. Article ID 573518.
[23] Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 2013:41.
[24] Caporaso J, Lauber C, Walters W, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 2012;6:21–4.
[25] Bacci G, Bazzicalupo M, Benedetti A, Mengoni A. StreamingTrim 1.0: a Java software for dynamic trimming of 16SrRNA sequence data from metagenetic studies. Mol Ecol Resour 2014;14:426–34.
[26] Bacci G, Ceccherini M, Bani A, Bazzicalupo M, Castaldini M, Galardini M, et al. Exploring the dynamics of bacterial community composition in soil: the pan-bacteriome approach. Antonie Van Leeuwenhoek 2015;107:785–97.
[27] Bacci G, Pagoto E, Passaponti M, Vannocci P, Ugolini A, Mengoni A. Composition of supralittoral sediments bacterial communities in a Mediterranean island. Ann Microbiol 2015;65:1–13.
[28] **Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. Vegan: community ecology package. R Package Version 2.0-10. 2013.** http://CRAN.R-project.org/package=vegan.
[29] Emiliani G, Mengoni A, Maida I, Perrin E, Chiellini C, Fondi M, et al. Linking bacterial endophytic communities to essential oils: clues from *Lavandula angustifolia* Mill. Evid Based Complement Alternat Med 2014;2014:16. Article ID 650905.
[30] Hardoim PR, Van Overbeek LS, Berg G, Pirttilä AM, Compant S, Campisano A, et al. The hidden world within plants: ecological and evolutionary considerations for defining functioning of microbial endophytes. Microbiol Mol Biol Rev 2015;79:293–320.
[31] Maida I, Chiellini C, Mengoni A, Bosi E, Firenzuoli F, Fondi M, Fani R. Antagonistic interactions between endophytic cultivable bacterial communities isolated from the medicinal plant *Echinacea purpurea*. Environ Microbiol 2015;18(8):2357–65.
[32] Mengoni A, Maida I, Chiellini C, Emiliani G, Mocali S, Fabiani A, et al. Antibiotic resistance differentiates *Echinacea purpurea* endophytic bacterial communities with respect to plant organs. Res Microbiol 2014.
[33] Gagne-Bourgue F, Aliferis KA, Seguin P, Rani M, Samson R, Jabaji S. Isolation and characterization of indigenous endophytic bacteria associated with leaves of switchgrass (*Panicum virgatum* L.) cultivars. J Appl Microbiol 2013;114:836–53.
[34] Ben Arfa A, Combes S, Preziosi-Belloy L, Gontard N, Chalier P. Antimicrobial activity of carvacrol related to its chemical structure. Lett Appl Microbiol 2006;43:149–54.
[35] Lopez-Romero JC, González-Ríos H, Borges A, Simões M. Antibacterial effects and mode of action of selected essential oils components against *Escherichia coli* and *Staphylococcus aureus*. Evid Based Complement Alternat Med 2015;2015:9.
[36] Villanueva Bermejo D, Angelov I, Vicente G, Stateva RP, García-Risco MR, Reglero G, et al. Extraction of thymol from different varieties of thyme plants using green solvents. J Sci Food Agric 2015;95:2901–7.
[37] Berg G, Krechel A, Ditz M, Sikora RA, Ulrichc A, Hallmann J. Endophytic and ectophytic potato-associated bacterial communities differ in structure and antagonistic function against plant pathogenic fungi. FEMS Microbiol Ecol 2005;51:215–29.
[38] Bai Y, Müller DB, Srinivas G, Garrido-Oter R, Potthoff E, Rott M, et al. Functional overlap of the *Arabidopsis* leaf and root microbiota. Nature 2015;528:364–9.
[39] Rosenblueth M, Martínez-Romero E. Bacterial endophytes and their interactions with hosts. Mol Plant-Microbe Interact 2006;19:827–37.

Please cite this article in press as: Checcucci A, et al., Is the plant-associated microbiota of *Thymus* spp. adapted to plant essential oil?, Research in Microbiology (2016), http://dx.doi.org/10.1016/j.resmic.2016.11.004

6.2 Exploiting nitrogen-fixing rhizobial symbionts genetic resources for improving phytoremediation of contaminated soils

This work is a review about the large variety of possible application of rhizobial symbionts to improve phytoremediation in contaminated areas.

Authors: **Alice Checcucci**, Marco Bazzicalupo, Alessio Mengoni

Department of Biology, University of Florence, via Madonna del Piano 6, 50019 Sesto Fiorentino (Firenze), Italy

in "Enhancing Cleanup of Environmental Pollutants: Biological Approaches", Editors Naser A. Anjum, Sarvajeet S. Gill, Narendra Tuteja,
Springer International Publishing AG, Cham. (2017). In Press.

6.2.1 Abstract

Rhizobia are one of the most relevant components of the plant-associated microbiota. They are found both in soil and associated, as commensals or symbionts with several plant taxa. In particular, with leguminous plants they establish a symbiotic association which allow the bacteria to express the enzyme nitrogenase responsible for the reduction of atmospheric dinitrogen. Consequently, rhizobia allow host plants to colonize marginal lands and nitrogen-deficient

soils, as for instance contaminated soils. The use of legume-rhizobial symbiosis for phytoremediation would allow to increase plant coverage (then

phytostabilization) of contaminated areas, without the need of expensive nitrogen fertizilization of the soil. Moreover, among host legumes, both pioneer plants (of for instance degraded lands) and crops (as alfalfa) are present, which allow an easy implementation of agronomical practices. Finally, the large genomic and phenotypic diversity of rhizobia allows the selection of élite strains resistant to harsh soil conditions and the creation of potentially new strains with the desired features for assisting legume-based phytoremediation.

6.2.2 The diversity of rhizobia

Rhizobia constitute a fraction of bacteria inhabiting on the root of plants (rhizobacteria), which have effect in promoting growth and alleviating the stress of the plant (plant growth-promoting rhizobacteria, PGPB) (Dimkpa et al., 2009;Lugtenberg & Kamilova, 2009). The PGPB may supply the host with a higher amount of nitrogen, synthesize several different phytohormones which can enhance various stages of plant growth, synthesize siderophores which can solubilise and sequester iron from the soil providing it to plant cells, have mechanisms for the solubilisation of phosphates, which then become more readily available, etc. (Lugtenberg & Kamilova, 2009). Rhizobia includes all those rhizobacteria which form endosymbiotic nitrogen fixing association with legumes. Rhizobia are a paraphyletic group of soil-inhabiting bacteria that fall into two classes of the *Proteobacteria*, the alpha- and beta-*proteobacteria* with the potential for establishing symbiotic

relationships with many legume plants. The first known species of rhizobia was *Rhizobium leguminosarum*, and with some subsequent known species was initially placed in Rhizobium genus (van Rhijn & Vanderleyden, 1995). Then, more advanced methods allow a re-classification into distinct genera (*Azorhizobium*, *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium* and *Sinorhizobium*/*Ensifer*).

Table 1. Genera and Species of the most known nitrogen-fixing PGPB and respective host legumes

Root-Nodule Bacteria		**Host Legume (genus)**
Genus	**Species**	
Azorhizobium	*caulidans*	*Sesbania*
Bradyrhizobium	*elkanii*	*Glycine*
//	*japonicum*	*Glycine*
//	*liaoningense*	*Glycine*
//	*yuanmingense*	*Lespedeza*
Mesorhizobium	*amorphae*	*Amorpha*
//	*chacoense*	*Prosopis*
//	*ciceri*	*Cicer*
//	*huakuii*	*Astragalus*
//	*loti*	*Lotus*
//	*mediterraneum*	*Cicer*
//	*plurifarium*	*Acacia , Leucaena*
//	*tianshanense*	*Glycyrrhiza , Sophora*
Rhizobium	*etli*	*Phaseolus*
//	*galegae*	*Galega , Leucanena*
	gallicum	*Phaseolus ,Dalea ,Onobrychis , Leucaena*
	giardinii	*Phaseolus*
	hainanense	*Stylosanthes , Centrosema*
	huautlense	*Sesbania*
	indigoferae	*Indigofera*
	leguminosarum bv trifolii	*Trifolium*
	leguminosarum bv viciae	*Pisum, Vicia , Lathyrus , Lens*

	leguminosarum bv phaseoli	*Phaseolus*
	loessense	*Astragalus*
	mongolense	*Medicago, Phaseolus*
	sullae	*Hedysarum*
	tropici	*Phaseolus, Leucaena, Dalea, Macroptilium*
	undicola	*Neptunia*
Sinorhizobium	*abri*	*Abrus*
	americanus	*Acacia*
	fredii	*Glycine*
	indianense	*Sesbania*
	kummerwiae	*Kunmerowia*
	medicae	*Medicago*
	meliloti	*Melilotus, Medicago, Trigonella*
	morelense	*Leucaena*
	saheli	*Sesbania*
	sahalense	*Sesbania*
	terangae	*Sesbania, Acacia*
	xinjiangnse	*Glycine*

The symbiotic relationship allows rhizobia to live inside the roots of the legume plant, consuming carbohydrates from the host and providing the legume with nitrogen that the bacteria convert into plant-usable form. Most of the symbiotic associations that involve legumes and rhizobia are characterized by a species-specificity, mainly due to the specific signal molecules present in the root exudates of legume plants that allow partners recognition and symbiosis development. In particular, rhizobia are able to sense specific secondary metabolites called flavonoids secreted by the roots of their host legume plant. The perception of the flavonoid signal modifies the rhizobial behaviour (Spini et al., 2015) and initiate a cascade of events which then lead to the entry of rhizobia inside plant root tissue and establish the symbiotic interaction

(Gibson et al., 2008). In fact, flavonoids trigger the secretion of Nod factors by the rhizobia that are recognized by transmembrane receptors on root hairs cells of specific legume: different strains of rhizobia produce different Nod factors, and different legumes produce receptors of different specificity. Following the root bacterial colonization, the root cortical cells begin to divide generating the nodules where the rhizobia differentiate into nitrogen fixing bacteroids. In effective nodules, the bacteria convert atmospheric nitrogen (N_2) into ammonia (NH_3), which then enter plant biosynthetic pathways via glutamine. This allow plants to grow better and thrive especially in nitrogen deficient soils.

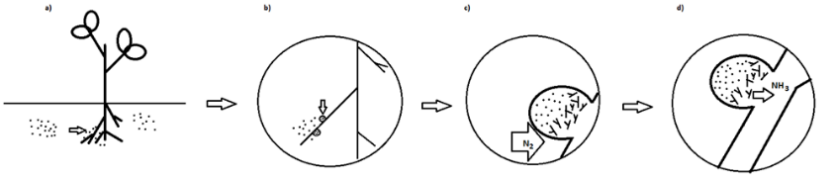


Figure 1. Steps of the symbiotic interaction. a) The symbiotic process starts with the approaching of the rhizobia present in the soil to the leguminous plant roots. b) The rhizobia are attracted by specific chemoattractants, the flavonoids, released by legume roots. Then the first step in developing the symbiosis is their consequent attachment to the surface of the plant root hair and the formation of the nodules. The nodules hold the bacterial symbiont providing specialized conditions necessary for nitrogen fixation. c) Inside the nodules part of the rhizobial population differentiate into nitrogen fixing bacteroids, d) the only form capable to transform molecular nitrogen in ammonia, thanks to the activity of the bacterial enzyme nitrogenase. The produced ammonia is then released within the plant.

This biological process is possible thanks to a multi-enzymatic complex called nitrogenase, active inside bacteroids. In return, the

rhizobia can benefit of the nutrients that the plant provides, and are protected inside the nodule structure. In ineffective nodules no or low nitrogen level is fixed but the rhizobia are still supplied with nutrients, and in this situation the bacteria could be considered parasitic (Denison & Kiers, 2004). To control the efficiency of the symbiotic partnership allowing the improvement of their growth, the plants have evolved mechanisms which favor nodules colonized by beneficial rhizobia (Denison, 2000).

6.2.3 Host legumes relevant for phytoremediation

Plant-associated bacteria have been recognized as one of the most relevant issues in improving phytoremediation yields (Glick, 2003;Afzal et al., 2014). In the last years, the association between leguminous plants and rhizobia has stirred the attention of researchers involved in the restoration of degraded lands and phytoremediation (Zahran, 1999;Sheaffer & Seguin, 2003;Wang et al., 2005;Hao et al., 2014;Teng et al., 2015). In particular, the possibility offered to cultivate legumes on marginal and nutrient-poor soils thanks to the intimate association with PGPB and in particular with nitrogen-fixing symbiotic partners has been seen as an opportunity to increase phytoremediation yields, while reducing its costs (Safronova et al., 2011; Hao et al., 2014). Among the most relevant legumes tried for phytoremediation are those belonging to genera *Lupinus* and *Sesbania* which have been used in the remediation of mine deposits. Interestingly, for the forage legume *M. sativa*, application trials for decontamination from PCB have been carried out, in symbiosis with *S. meliloti* strains, possibly able to degrade such recalcitrant compounds. In some cases, interesting

results were obtained when the symbiotic partnership between host legume and rhizobial symbiont included also arbuscular mycorrhizal fungi, which allowed to reduce the toxic effect of the contaminant (Garg & Bhandari, 2012)

Table 2. Some legumes used in phytoremediation*.

Host legume	**Use in phytoremediation**	**Rhizobial symbiont**	**Reference**
Anthyllis vulneraria	Heavy metals	*Mesorhizobium metallidurans*	Vidal et al, 2009
Cajanus cajan	Cd	*Rhizobium sp.*	Garg and Bhandari , 2012
Cicer arietinum	Cr	*Mesorhizobium sp.*	Wani et al., 2008
Cicer arietinum	Zn, Cu, Cr, Cd, Fe	*Rhizobium sp.*	Gupta et al., 2004
Glycine max	Cd	*Pseudomonas putida, Pseudomonas monteilli*	Rani et al., 2009
Lablab purpureus	Spent engine oil, Cu	*Rhizobium sp.,Bradyrhizobium lablabi*	Younis, 2007; Ismail et al., 2014
Lens culinaris	Zn	*Rhizobium leguminosarum*	Wani et al., 2008
Leucaena	Cd, Zn	*Rhizobium sp.*	Saraswat and Rai, 2011
Lolium multiflorum	Cd, Zn and Ni	*Bradyrhizobium sp.*	Wani, 2007; Guo, 2014
Lotus edulis	Cd, Pb and Zn	*Mesorhizobium loti*	Safronova, 2010
Lotus ornithopodioides	Cd, Pb and Zn	*Mesorhizobium loti*	Safronova, 2010
Lupinus albus	Cd, Cu, Pb and Zn	*Bradyrhizobium sp.,*	Pajuelo et al., 2008

		Ochrobactrum sp	
Lupinus luteus	Cd, Cu, Pb and Zn	*Bradyrhizobium sp., Ochrobactrum sp*	Pajuelo et al., 2008
Medicago ciliaris	Cd, Pb and Zn	*Sinorhizobium sp.*	Safronova, 2010
Medicago sativa	As	*Sinorhizobium sp.*	Pajuelo et al., 2008
Medicago sativa	polychlorinated biphenyls (PCB)	*S. meliloti*	Mehmannavaz a, 2002
Mimosa pudica	Pb, Cu and Cd	*Cupravirus taiwanensis*	Chen, 2008
Pisum sativum	Ni and Zn	*Rhizobium sp.*	Wani et al., 2008
Pisum sativum	Cd	*Rhizobium leguminosarum*	Engqvist et al., 2006
Prosopis juliflora	Fe, Mn, Cu, Zn, Cr and Pb	*Rhizobium sp.*	Rai et al.,2004 Sinha et al., 2005
Robinia pseudoacacia	Zn	*Agrobacterium tumefaciens*	Smith, 1992
Robinia pseudocacia	Cu	*Mesorhizobium amorphae*	Hao et al., 2015
Sesbania cannabina	Pb and Zn	*Azorhizobium caulinodans*	Chan et al., 2003
Sesbania grandiflora	Pb and Zn	*Azorhizobium caulinodans*	Chan et al., 2003
Sesbania rostrata	Pb and Zn	*Azorhizobium caulinodans*	Chan et al., 2003 Shuguang et al., 2009
Sesbania sesban	Pb and Zn	*Azorhizobium caulinodans*	Chan et al., 2003
Vigna mungo	Cd	*Pseudomonas aeruginosa strains*	Ganesan, 2008
Vigna radiata	Cd, Zn and Ni	*Bradyrhizobium sp.*	Wani, 2007 Guo, 2014

Vigna radiata	Cr	*Ochromobactrum intermedium*	Faisal and Hasnain, 2006

* Host legume, rhizobial symbiont and the related use in phytoremediation are reported.

6.2.4 Use of rhizobia in phytoremediation

Tough potentially tolerant to pollutants, legumes relevant for phytoremediation can be limited in their growth. Symbiotic rhizobia, providing fixed nitrogen can allow to promote host legumes growth in marginal and degraded soils, as usually are those claimed for restoration. Several efforts have been carried out both on improvement of plant traits to cope with unfavourable conditions (Dwivedi et al., 2010;Dwivedi et al., 2015). Since the efficiency of the process of N2 fixation is related to both the physiological state of the host plant and to rhizobial partner (Zahran, 1999), efforts have been made to find symbiotic rhizobia which can tolerate harsh conditions. Rhizobial strains in natural environments showing tolerance to various stressors (e.g. salinity, heavy-metals, pH, etc.) have been found and tested as inocula on target plant (Hungria et al., 1993;Zahran, 2001;Provorov & Tikhonovich, 2003;Roumiantseva, 2009; Elboutahiri et al., 2010;Trabelsi et al., 2010;Boukhatem et al., 2012).

Concerning salt stress, this inhibits the initial steps of symbiosis (Coba de la Peña et al., 2003). Several studies have been performed looking at salt tolerant rhizobia, especially in sub-arid regions, where conditions may likely have contributed in selecting rhizobial strains with the ability to cope with osmotic stress (Mnasri et al., 2007;Trabelsi et al., 2010) These studies showed the presence of

rhizobia tolerant to high NaCl concentrations (up to 1M), which can be used as inocula for crop production in salinized soils. Often linked to salt stress are also drought and temperature stresses (Alexandre & Oliveira, 2013). In particular, desiccation is a critical step if proper inocula have to be prepared for spraying as biofertilizers as well as for long term survival of inoculated rhizobia in arid soils. An adaptation to desiccation has been shown for strains which experienced salt stress before desiccation, but the molecular mechanisms of desiccation survival, which may involve bet genes, as well as exopolysaccharide production, are still not fully understood (Vriezen et al., 2007).

Related to the phytoremediation, highly relevant is the contamination by heavy-metals. In last few years there has been an increasing interest in rhizobial symbionts from wild legumes growing in soil rich in trace metals (as nickel, copper, etc.). In particular the flora of serpentine soils has been deeply investigated. Serpentine soils are distributed all over the world and originate from an array of ultramafic rock types characterized by high levels of nickel, cobalt and chromium, low levels of N, P, K, Ca, and a high Mg/Ca ratio (Brooks, 1987). The flora of serpentine soil contains several endemics, including many legume species (Brady et al., 2005) The bacteria inhabiting serpentine soil and endophytes of serpentine plants have attracted the attention of many investigators (Mengoni et al., 2010) and references therein). In particular, several bradyrhizobial strains with tolerance up to 15 mM Ni(II) have been isolated from the endemic legume *Serianthes calycina* growing in New Caledonia serpentine soils (Chaintreuil et al., 2007). Moreover, metal tolerant rhizobia have been isolated also from legume species growing on mine deposits. A new highly tolerant to Zn rhizobial

species (*Mesorhizobium metallidurans*) has been identified as symbiont of *Anthyllis vulneraria*, a legume species growing in zinc mine deposits (Vidal, C et al., 2009). Interestingly, the association between *A. vulneraria* and *M. metallidurans* has been demonstrated effective for the growth of the host plant in soil contaminated by Zn, Pb and Cd (Mahieu et al., 2011). *A. vulneraria / M. metallidurans* association was able to promote *A. vulneraria* growth on strongly contaminated soil (16% Zn, 9.2% Pb and 1.3% Cd) and to obtain fix up to 80% of its total nitrogen from atmospheric N_2. In conclusion, there are several rhizobial strains which have been isolated and characterized for tolerance to many environmental stresses and proved to be effective in improving legume growth under unfavourable conditions as for instance as pioneer species for restoration ecology in marginal lands (Wang et al., 2005).

6.2.5 Genomics as a way for improving rhizobial performances

Rhizobial tolerance to pollutants is related to the presence of specific genes or metabolic pathways able to degrade the pollutant (e.g. organic pollutants) or tolerate the toxic effects (e.g. trace metals). Consequently, to improve phytoremediation yields of rhizobial-legume partnerships, selection of tolerant élite rhizobial strains (Dwivedi et al., 2015) is as crucial as is that of tolerant plant germplams. In the last years, thanks to the efforts of the scientific communities in the sequencing of the genomes of a plethora rhizobacteria and rhizobia (Reeve et al., 2015), a lot of information is available which can be used to identify, by genome analysis, useful strains or gene traits. In particular, comparative genome

analysis has helped to identify additional genetic determinants for plant-bacteria interaction relevant phenotypes (Mengoni et al., 2014;Galardini et al., 2015). In fact, bacterial genomes are composed by two main part, a common gene set (shared by all members of a given taxon) named core genome and by a dispensable gene set (dispensable genome) which is specific of some members only (Tettelin et al., 2005;Tettelin et al., 2008). Core genome includes genes which confer the taxon identity (e.g. basic cellular machinery, housekeeping metabolism, etc.), while the dispensable genome fraction is related to genes which confer strain-specific features, such as environmental adaptation. The modular feature of the bacterial pangenome (as sum of core and dispensable genome fractions), allow bacterial strains to easily adapt to changing environmental conditions by integrating genes (and replicons, such as accessory plasmids) into the core genome machinery (Young et al., 2006). It is consequently plausible that genes related to improving plant-bacteria relationships in terms of plant productivity and especially plant tolerance to abiotic stress, can be found in the dispensable genome fraction. The analysis of rhizobial pangenomes (viz. the comparative genome analysis of conspecific strains) can then disclose the genetic basis of their different tolerance to pollutants allowing to potentially design new strains which may combine the genetic relevant features of different individual strains (e.g. for both high nitrogen fixation rate and heavy-metal tolerance). The modular nature (diCenzo et al., 2014) and redundant feature (González et al., 2006;diCenzo & Finan, 2015) of many rhizobial genomes helps in defining nearly-independent genetic cassettes (even containing several genes), which can be delivered through the large plasmids typical of many rhizobial strains in a sort of

"pangenome-assisted strain improvement" or in analogy with plant breeders.

For the alfalfa symbiont *Sinorhizobium meliloti*, the pangenome analysis approach has allowed to identify genes related to the different cooperative behaviour of strains (Galardini et al., 2011;Galardini et al., 2013), which could be used for improving the symbiotic performance of strains. Moreover, the same analysis has helped to identify and characterize a core genome determinant (Sma1641), encoding an ortholog of nreB gene involved in Ni^{+} efflux (Grass et al., 2001). This gene has been shown to have pleiotropic effects also toward the tolerance to Cu^{+} (Pini et al., 2013) and in particular, its deletion, induced a higher plant growth under in vitro culture conditions (Pini et al., 2013), indicating that metal homeostasis, symbiosis and plant growth promotion can be tightly linked. A copper tolerant *S. meliloti* strain (CCNWSX0020) has been isolated from mine tailing and its genome has been completely sequenced (Li et al., 2012) and shown to positively influence the growth of the host legume *Medicago lupulina* in copper contaminated soils (Li et al., 2014). Functional genomics analyses of *S. meliloti* CCNWSX0020 revealed that genes involved in copper homeostasis (copper chelation and export to the bacterial periplasm) were responsible for the copper tolerance (Li et al., 2014). Interestingly, those genes are widespread in Rhizobiaceae and several genomes of strains of *S. meliloti*, as well as *R. leguminosarum,* contain such genes (Figure 2), indicating that common cellular mechanisms (then being part of the core genome) can be used also, in adjunct to dispensable genes, for biotechnological improvement of élite strains. A similar investigation carried out on a copper tolerant strain of

Mesorhizobium amorphae nodulating *Robinia pseudoacacia* indicated that a P-type ATPase, possibly involved in the efflux of cytoplasmic copper can be one of the determinants of the copper tolerant phenotype (Hao et al., 2015). However, the same authors found that other genes, with unknown function are related to copper tolerance in such strain. This latter evidence indicate that our understanding of heavy-metal tolerance in bacteria is still far from being complete.

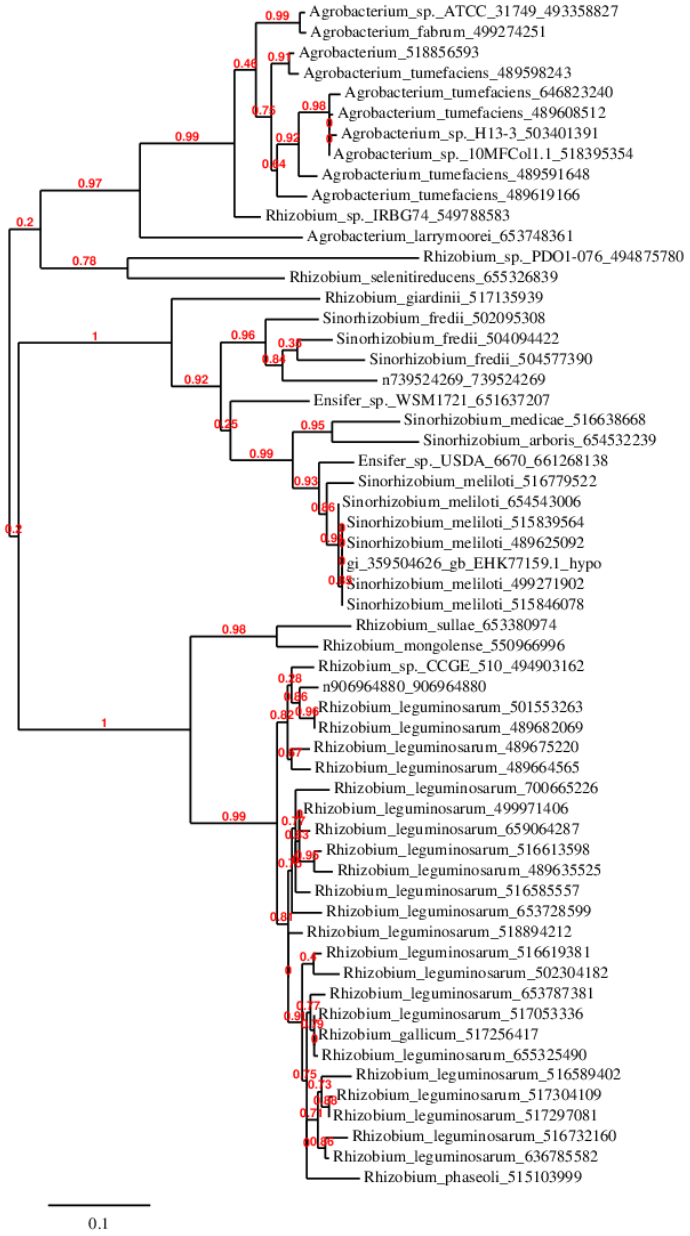


Figure 2. Widespread occurrence of genes for metal homeostasis in rhizobia. Maximum likelihood phylogenetic tree of the orthologs of the protein SM0020_14779 from *S. meliloti* CCNWSX0020. Number at nodes indicate bootstrap values. Arrow indicates the protein SM0020_14779 (gi|359504626|gb|EHK77159.1|).

The genomic features of rhizobia in terms of modular and functionally redundant genomes and the large pangenomes they harbour, allow the search for both élite strains and élite genes to be used in phytoremediation. In particular, élite genes (as for instance the above mentioned Sma1641 from *S. meliloti*) could be transferred from tolerant to sensitive strains to possibly combine tolerant features with strain competitiveness and other important features for the symbiotic interaction (Dwivedi et al., 2015). This perspective of rhizobial synthetic biology has granted credits and genome sequencing of relevant strains have been promoted by the USA Department of Energy – Joint Genome Institute have been performed (Biondi et al., 2009).

6.2.6 Conclusions

We have presented here a brief summary of the use of the legume-rhizobium symbiosis for increasing phytoremediation yields and application. There is now a number of investigation, mainly done under laboratory and controlled conditions, which indicate that rhizobia with the ability to tolerate toxic compound and/or to degrade contaminants are present in nature. These bacteria can be found as symbionts of legumes thriving on contaminated soils, by either geological features (e.g. serpentine outcrops) and anthropic causes (e.g. mine deposits). The study of their physiology and of their genome may enable to use them (or their genes) as partner of the same host legumes, or of legumes for which better growth and agronomic potential are present. In this respect, the use of legume crops in phytoremediation could allow the use of already settled

agronomic practices, then reducing the cost and time of the translation of lab-based evidences to the field

Bibliography

Afzal M, Khan QM, Sessitsch A. 2014. Endophytic bacteria: Prospects and applications for the phytoremediation of organic pollutants. Chemosphere 117(1): 232-242.

Alexandre A, Oliveira S. 2013. Response to temperature stress in rhizobia. Critical Reviews in Microbiology 39(3): 219-228.

Biondi EG, Mengoni A, Brilli M, Bazzicalupo M, Mocali S. 2009. Systematic approach for the improvement of alfalfa production using genomic diversity of *Sinorhizobium meliloti* natural populations. 2009 JGI User Meeting: 1.

Boukhatem ZF, Domergue O, Bekki A, Merabet C, Sekkour S, Bouazza F, Duponnois R, de Lajudie P, Galiana A. 2012. Symbiotic characterization and diversity of rhizobia associated with native and introduced acacias in arid and semi-arid regions in algeria. Fems Microbiology Ecology 80(3): 534-547.

Brady KU, Kruckeberg AR, Bradshaw HD. 2005. Evolutionary ecology of plant adaptation to serpentine soils. Annual Review of Ecology Evolution and Systematics 36: 243-266.

Brooks RR. 1987. Serpentine and its vegetation. A multidisciplinary approach. 9999 S.W. Wilshire, Portland, OR 97225: Dioscorides Press.

Chaintreuil C, Rigault F, Moulin L, Jaffre T, Fardoux J, Giraud E, Dreyfus B, Bailly X. 2007. Nickel resistance determinants in *Bradyrhizobium* strains from nodules of the endemic new caledonia legume serianthes calycina. Appl Environ Microbiol 73(24): 8018-8022.

Chan GYS, Zhi HY, Wong MH. 2003. Comparison of four sesbania species to remediate pb/zn and cu mine tailings. Environmental management 32(2): 246-251.

Chen WM, Wu,C.H.,James,E.K.,andChang,J.S. 2008. Biosorption capability of cupriavidus taiwanensis and its effects on heavy metal removal by nodulated mimosa pudica. J. Hazard.Mater 151: 364–371.

Coba de la Peña T, Verdoy D, Redondo F, Pueyo JJ. 2003. Salt tolerance in the rhizobium-legume symbiosis: An overview. Plant Mol. Biol 1: 187-205.

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

Denison RF, Kiers ET. 2004. Why are most rhizobia beneficial to their plant hosts, rather than parasitic? Microbes and Infection 6(13): 1235-1239.

diCenzo GC, Finan TM. 2015. Genetic redundancy is prevalent within the 6.7 mb *Sinorhizobium meliloti* genome. Molecular Genetics and Genomics: 1-12.

diCenzo GC, MacLean AM, Milunovic B, Golding GB, Finan TM. 2014. Examination of prokaryotic multipartite genome evolution through experimental genome reduction. PLoS Genet 10(10): e1004742.

Dimkpa C, Weinand T, Asch F. 2009. Plant–rhizobacteria interactions alleviate abiotic stress conditions. Plant, Cell & Environment 32(12): 1682-1694.

Dwivedi SL, Sahrawat KL, Upadhyaya HD, Mengoni A, Galardini M, Bazzicalupo M, Biondi EG, Hungria M, Kaschuk G, Blair MW, Ortiz R 2015. Advances in host plant and rhizobium genomics to enhance symbiotic nitrogen fixation in grain legumes.In. Advances in Agronomy. 1-116.

Dwivedi SL, Upadhyaya HD, Balaji J, Buhariwalla HK, Blair MW, Ortiz R, Crouch JH, Serraj R 2010. Using genomics to exploit grain legume biodiversity in crop improvement. Plant breeding reviews: John Wiley & Sons, Inc., 171-357.

Elboutahiri N, Thami-Alami I, Udupa S. 2010. Phenotypic and genetic diversity in sinorhizobium meliloti and s. Medicae from drought and salt affected regions of morocco. BMC Microbiology 10(1): 15.

Engqvist LG, MÃ¥rtensson A, Orlowska E, Turnau K, Belimov AA, Borisov AY, Gianinazzi-Pearson V. 2006. For a successful pea production on polluted soils, inoculation with beneficial microbes requires active interaction between the microbial components and the plant. Acta Agriculturae Scandinavica Section B-Soil and Plant Science 56(1): 9-16.

Faisal M, Hasnain S. 2006. Growth stimulatory effect of ochrobactrum intermedium and bacillus cereus on *Vigna radiata* plants. Letters in applied microbiology 43(4): 461-466.

Galardini M, Bazzicalupo M, Biondi E, Brambilla E, Brilli M, Bruce D, Chain P, Chen A, Daligault H, Walston Davenport K, Deshpande S, Detter JC, Goodwin LA, Han C, Han J, Huntemann M, Ivanova N, Klenk H-P, Kyrpides NC, Markowitz V, Mavrommatis K, Mocali S, Nolan M, Pagani I, Pati A, Pini F, Pitluck S, Spini G, Szeto E, Teshima H, Woyke T, Mengoni A. 2013. Permanent draft genome sequences of the symbiotic nitrogen fixing *Ensifer meliloti* strains bo21cc and ak58. Standards in Genomic Sciences 9: 2.

Galardini M, Mengoni A, Brilli M, Pini F, Fioravanti A, Lucas S, Lapidus A, Cheng J-F, Goodwin L, Pitluck S, Land M, Hauser L, Woyke 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(1): 235.

Galardini M, Mengoni A, Mocali S 2015. From pangenome to panphenome and back. In: A. Mengoni, M. Fondi, M. Galardini eds. Bacterial pangenomics: Springer New York, 257-270.

Ganesan V. 2008. Rhizoremediation of cadmium soil using a cadmium-resistant plant growth-promoting rhizopseudomonad. Current microbiology 56(4): 403-407.

Garg N, Bhandari P. 2012. Influence of cadmium stress and arbuscular mycorrhizal fungi on nodule senescence in *Cajanus cajan* (l.) millsp. International Journal of Phytoremediation 14(1): 62-74.

Gibson KE, Kobayashi H, Walker GC. 2008. Molecular determinants of a symbiotic chronic infection. Annual Review of Genetics 42: 413-441.

Glick BR. 2003. Phytoremediation: Synergistic use of plants and bacteria to clean up the environment. Biotechnology Advances 21(5): 383-393.

González V, Santamaría RI, Bustos P, Hernández-González I, Medrano-Soto A, Moreno-Hagelsieb G, Janga SC, Ramírez MA, Jiménez-Jacinto V, Collado-Vides J. 2006. The partitioned *Rhizobium etli* genome: Genetic and metabolic redundancy in seven interacting replicons. Proceedings of the National Academy of Sciences of the United States of America 103(10): 3834-3839.

Grass G, Fan B, Rosen BP, Lemke K, Schlegel HG, Rensing C. 2001. Nreb from *Achromobacter xylosoxidans* 31a is a nickel-induced transporter conferring nickel resistance. J Bacteriol 183(9): 2803-2807.

Guo JK, andChi,J. 2014. Effect of cd-tolerant plant growth-promoting rhizobium on plant growth and cd uptake by *Lolium multiflorum* lam. and glycinemax (l.) merr.In cd-contaminated soil. PlantSoil 375: 205-214.

Gupta DK, Rai UN, Sinha S, Tripathi RD, Nautiyal BD, Rai P, Inouhe M. 2004. Role of rhizobium (ca-1) inoculation in increasing growth and metal accumulation in cicer arietinum l. Growing under fly-ash stress condition. Bulletin of environmental contamination and toxicology 73(2): 424-431.

Hao X, Taghavi S, Xie P, Orbach MJ, Alwathnani HA, Rensing C, Wei G. 2014. Phytoremediation of heavy and transition metals aided by legume-rhizobia symbiosis. International Journal of Phytoremediation 16(2): 179-202.

Hao X, Xie P, Zhu Y-G, Taghavi S, Wei G, Rensing C. 2015. Copper tolerance mechanisms of *Mesorhizobium amorphae* and its role in aiding phytostabilization by robinia pseudoacacia in copper contaminated soil. Environmental science & technology 49(4): 2328-2340.

Hungria M, Franco AA, Sprent JI. 1993. New sources of high-temperature tolerant rhizobia for *Phaseolus vulgaris* l. Plant and Soil 149(1): 103-109.

Ismail H, Ijah U, Riskuwa M, Allamin I. 2014. Biodegradation of spent engine oil by bacteria isolated from the rhizosphere of legumes grown in contaminated soil. International Journal of Environment 3(2): 63-75.

Li Z, Ma Z, Hao X, Rensing C, Wei G. 2014. Genes conferring copper resistance in *Sinorhizobium meliloti* ccnwsx0020 also promote the growth of *Medicago lupulina* in copper-contaminated soil. Applied and Environmental Microbiology 80(6): 1961-1971.

Li Z, Ma Z, Hao X, Wei G. 2012. Draft genome sequence of *Sinorhizobium meliloti* ccnwsx0020, a nitrogen-fixing symbiont with copper tolerance capability isolated from lead-zinc mine tailings. Journal of Bacteriology 194(5): 1267-1268.

Lugtenberg B, Kamilova F. 2009. Plant-growth-promoting rhizobacteria. Annual Review of Microbiology 63(1): 541-556.

Mahieu S, Frèrot H, Vidal Cl, Galiana A, Heulin K, Maure L, Brunel B, Lefébvre C, Escarré J, Cleyet-Marel J-C. 2011. *Anthyllis vulneraria/mesorhizobium metallidurans*, an efficient symbiotic nitrogen fixing association able to grow in mine tailings highly contaminated by zn, pb and cd. Plant and Soil 342(1-2): 405-417.

Mehmannavaza R, Prasher, S.O., and Ahmad, D. 2002. Rhizospherice ffects of alfalfaonbiotransformationofpolychlorinatedbiphenylsinacontaminated soil augmented with *Sinorhizobium meliloti*. ProcessBiochem 35: 955–963.

Mengoni A, Galardini M, Fondi M. 2014. Bacterial pangenomics: Methods and protocols: Springer New York.

Mengoni A, Schat H, Vangronsveld J. 2010. Plants as extreme environments? Ni-resistant bacteria and ni-hyperaccumulators of serpentine flora. Plant and Soil 331(1): 5-16.

Mnasri B, Mrabet M, Laguerre G, Aouani M, Mhamdi R. 2007. Salt-tolerant rhizobia isolated from a tunisian oasis that are highly effective for symbiotic n2-fixation with phaseolus vulgaris constitute a novel biovar (bv. Mediterranense) of *Sinorhizobium meliloti*. Archives of Microbiology 187(1): 79-85.

Pajuelo E, Dary M, Palomares AJ, Rodriguez-Llorente ID, Carrasco JA, Chamber MA 2008. Biorhizoremediation of heavy metals toxicity using rhizobium-legume symbiosis. Biological nitrogen fixation: Towards poverty alleviation through sustainable agriculture: Springer, 101-104.

Pini F, Spini G, Galardini M, Bazzicalupo M, Benedetti A, Chiancianesi M, Florio A, Lagomarsino A, Migliore M, Mocali S, Mengoni A. 2013. Molecular phylogeny of the nickel-resistance gene nreb and functional role in the nickel sensitive symbiotic nitrogen fixing bacterium *Sinorhizobium meliloti*. Plant and Soil.

Provorov NA, Tikhonovich IA. 2003. Genetic resources for improving nitrogen fixation in legume-rhizobia symbiosis. Genetic Resources and Crop Evolution 50(1): 89-99.

Rai UN, Pandey K, Sinha S, Singh A, Saxena R, Gupta DK. 2004. Revegetating fly ash landfills with *Prosopis juliflora* l.:

Impact of different amendments and rhizobium inoculation. Environment International 30(3): 293-300.

Rani A, Souche YS, Goel R. 2009. Comparative assessment of in situ bioremediation potential of cadmium resistant acidophilic *Pseudomonas putida* 62bn and alkalophilic *Pseudomonas monteilli* 97an strains on soybean. International Biodeterioration & Biodegradation 63(1): 62-66.

Reeve W, Ardley J, Tian R, Eshragi L, Yoon JW, Ngamwisetkun P, Seshadri R, Ivanova NN, Kyrpides NC. 2015. A genomic encyclopedia of the root nodule bacteria: Assessing genetic diversity through a systematic biogeographic survey. Stand Genomic Sci 10: 14.

Roumiantseva ML. 2009. Genetic resources of nodule bacteria. Russian Journal of Genetics 45(9): 1013-1026.

Safronova V, Piluzza G, Bullitta S, Belimov A. 2011. Use of legume-microbe symbioses for phytoremediation of heavy metal polluted soils: Advantages and potential problems. Handbook of Phytoremediation. Nova Science Publishers, Inc: 443-469.

Safronova VP, G; Belimov, AA; Bullitta S. 2010. Combined inoculation with rhizosphere and nodule bacteria improves growth, nutrition and heavy metal uptake of pasture legumes grown in polluted mine waste. Biol. Fertil. Soil (BFS0-D-10-00240, under review).

Saraswat S, Rai JPN. 2011. Prospective application of *Leucaena leucocephala* for phytoextraction of cd and zn and nitrogen fixation in metal polluted soils. International Journal of Phytoremediation 13(3): 271-288.

Sheaffer CC, Seguin P. 2003. Forage legumes for sustainable cropping systems. Journal of Crop Production 8(1-2): 187-216.

Shuguang J, Weijun S, Zhongyi Y. 2009. Enhanced adaptability of *Sesbania rostrata* to pb/zn tailings via stem nodulation. Journal of Environmental Sciences 21(8): 1135-1141.

Sinha S, Rai UN, Bhatt K, Pandey K, Gupta AK. 2005. Fly-ash-induced oxidative stress and tolerance in prosopis juliflora l. Grown on different amended substrates. Environmental monitoring and assessment 102(1-3): 447-457.

Smith SR, and Giller,K.E. 1992. Effective rhizobium leguminosarum bio-var trifolii present in five soils contaminated with heavy metals from long-term applications of sewages ludge or metalminespoil. Soil Biol.Biochem. 24: 781-788.

Spini G, Decorosi F, Cerboneschi M, Tegli S, Mengoni A, Viti C, Giovannetti L. 2015. Effect of the plant flavonoid luteolin on *Ensifer meliloti* 3001 phenotypic responses. Plant and Soil: 1-20.

Teng Y, Wang X, Li L, Li Z, Luo Y. 2015. Rhizobia and their bio-partners as novel drivers for functional remediation in contaminated soils. Frontiers in Plant Science 6.

Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin A. 2005. Genome analysis of multiple pathogenic isolates of streptococcus agalactiae: Implications for the microbial "Pan-genome". Proc Natl Acad Sci USA 102: 13950 - 13955.

Tettelin H, Riley D, Cattuto C, Medini D. 2008. Comparative genomics: The bacterial pan-genome. Current Opinion in Microbiology 11(5): 472-477.

Trabelsi D, Mengoni A, Aouani ME, Bazzicalupo M, Mhamdi R. 2010. Genetic diversity and salt tolerance of *Sinorhizobium* populations from two tunisian soils. Annals of Microbiology: 1-7.

Van Rhijn P, Vanderleyden J. 1995. The rhizobium-plant symbiosis. Microbiol Rev 59(1): 124-142.

Vidal Cl, Chantreuil Cm, Berge O, Mauré L, Escarré J, Béna G, Brunel B, Cleyet-Marel J-C. 2009. *Mesorhizobium metallidurans* sp. Nov., a metal-resistant symbiont of *Anthyllis vulneraria* growing on metallicolous soil in languedoc, france. International Journal of Systematic and Evolutionary Microbiology 59(4): 850-855.

Vriezen JAC, de Bruijn FJ, Nüsslein K. 2007. Responses of rhizobia to desiccation in relation to osmotic stress, oxygen, and temperature. Appl Environ Microbiol 73: 3451–3459.

Wang HB, Shu WS, Lan CY. 2005. Ecology for heavy metal pollution: Recent advances and future prospects. Acta Ecologica Sinica 25(3): 596-605.

Wani PA, Khan MS, Zaidi A. 2008a. Chromium-reducing and plant growth-promoting *Mesorhizobium* improves chickpea

growth in chromium-amended soil. Biotechnology letters 30(1): 159-163.

Wani PA, Khan MS, Zaidi A. 2008b. Effect of metal-tolerant plant growth-promoting rhizobium on the performance of pea grown in metal-amended soil. Archives of environmental contamination and toxicology 55(1): 33-42.

Wani PA, Khan MS, Zaidi A. 2008c. Impact of zinc-tolerant plant growth-promoting rhizobacteria on lentil grown in zinc-amended soil. Agronomy for sustainable development 28(3): 449-455.

Wani PA, Khan,M.S.,andZaidi,A. 2007. Effect of metal tolerant plant growth promoting *Bradyrhizobium* sp.(vigna) on growth, symbiosis, seedyield, and metal uptake by greengram plants. Chemosphere 70: 36-45.

Young JP, Crossman L, Johnston A, Thomson N, Ghazoui Z, Hull K, Wexler M, Curson A, Todd J, Poole P, Mauchline T, East A, Quail M, Churcher C, Arrowsmith C, Cherevach I, Chillingworth T, Clarke K, Cronin A, Davis P, Fraser A, Hance Z, Hauser H, Jagels K, Moule S, Mungall K, Norbertczak H, Rabbinowitsch E, Sanders M, Simmonds M, Whitehead S, Parkhill J. 2006. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. Genome Biology 7(4): R34.

Younis M. 2007. Responses of lablab purpureus-rhizobium symbiosis to heavy metals in pot and field experiments. World Journal of Agricultural Sciences 3(1): 111-122.

Zahran H. 2001. Rhizobia from wild legumes: Diversity, taxonomy, ecology, nitrogen fixation and biotechnology. 91(2-3): 143-153.

Zahran HH. 1999. Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. Microbiol Mol Biol Rev 63(4): 968–989.

6.3 The Integrated Microbial Genome resource of analysis

This part includes a contribute to the book “Bacterial Pangenomics: Methods and Protocol, Methods in Molecular Biology” vol 1231, doi:10.1007/978-1-4939-1720-4_18, Eds Alessio Mengoni et al., Springer Science+Buisness Media, New York, and describes the pipeline of genomic comparative analysis by the use of IMG tool. This tool was used during the preparation of some analyses reported in the present thesis, as the ACC deaminase phylogenyetic reconstruction.

Chapter 18

The Integrated Microbial Genome Resource of Analysis

Alice Checcucci and Alessio Mengoni

Abstract

Integrated Microbial Genomes and Metagenomes (IMG) is a biocomputational system that allows to provide information and support for annotation and comparative analysis of microbial genomes and metagenomes. IMG has been developed by the US Department of Energy (DOE)-Joint Genome Institute (JGI). IMG platform contains both draft and complete genomes, sequenced by Joint Genome Institute and other public and available genomes. Genomes of strains belonging to Archaea, Bacteria, and Eukarya domains are present as well as those of viruses and plasmids. Here, we provide some essential features of IMG system and case study for pangenome analysis.

Key words Genome database, Metagenome database, Integrated Microbial Genomes and Metagenomes, Joint Genome Institute, Bioinformatics, Genome comparison

1 Introduction

Integrated Microbial Genomes and Metagenomes (IMG, URL: https://img.jgi.doe.gov) is a biocomputational system that allows to provide information and support for annotation and comparative analysis of microbial genomes and metagenomes [1, 2]. IMG has been developed by the US Department of Energy (DOE)-Joint Genome Institute (JGI) and is one of the JGI database resources belonging to its Genome Portal (http://genome.jgi.doe.gov). The Genome Portal has a "Tree of Life" data organization, where the sequenced genomes are arranged by domains (and kingdom, phylum, class, or order) and metagenomes by the niche.

IMG platform contains both draft and complete genomes, sequenced by Joint Genome Institute and other public and available genomes. Genomes of strains belonging to Archaea, Bacteria, and Eukarya domains are present as well as those of viruses and plasmids. On December 31, 2013, IMG stored more than 18,000 genomes, 13,334 of which are bacterial.

The genome storage can be investigated by comparisons on single or multiple genes, at the genome scale and by single or

Alessio Mengoni et al. (eds.), *Bacterial Pangenomics: Methods and Protocols*, Methods in Molecular Biology, vol. 1231, DOI 10.1007/978-1-4939-1720-4_18,

multiple functions. The system is therefore composed by three kinds of genome analysis: primary sequences, genic model (annotation) and biocomputational predictions, and functional connection and pathway information.

Finally, IMG provides users some linked tools to support comparative microbial genes, genomes, and metagenomics analysis, including COG, KEGG, Pfam, InterPro, and the Gene Ontology. Consequently, thanks also to the graphical user interface IMG is particularly suited for **nonexperienced bioinformaticians** which want to perform comparative genome analyses.

The two main functions available in the platform are

- **Exploration of data**
- **Genome comparison analyses**

2 Exploring Data on IMG

With the buttons "Find", it is possible to start the data scanning of genomes, genes, and then functions and metabolic pathways according to various biocomputational tools, as Blast, COG, KOG, Pfam, TIGRfam, and KEGG.

2.1 Find Genomes

One of the most used functions is the **Genome Browser**, where all the genomes filed in the platform are listed *alphabetically* or as *phylogenetic tree*. Every genome is described for domain, status, study name, sequencing center, size of the genome, and number of genes found.

The **Genome Search** function is instead used to search for that particular genome which is of interest of the user. Search filters as the simple name or metadata ("data about data") categories. Concerning metadata values as phenotype, habitat, disease, relevance, geographic location and host can be used. *Individual genomes* can be examined with *organism details* page that can be accessed by clicking on a genome name in every list of genomes collected in IMG.

The information page of every organism contains four main sections:

- The *Overview* that includes genome information as sequencing, taxonomic classification, metadata, metabolism, publications, and NCBI ID.
- The *Genome statistics* provides information about DNA sequence, as GC content, annotation, scaffold, and cluster gene according to the main tools (COG, KOG, Pfam, TIGRfam).
- The *Viewer* section shows linear or circular chromosome map of the organism, its scaffolds, and contigs.
- The *Export* section allows to move and save genome sequence or data in a variety of formats detectable using Excel.

Other interesting functions collect specific data of the genome, as *phylogenetic distribution of genes*, that allow to observe the distribution of the genes using Blast on the IMG dataset genome, or *horizontally transferred genes*, that gives statistics about gene or sequence moving during the evolution.

2.2 Find Genes

With this function it is possible to search a single gene or a group of genes (as for instance an entire operon) in selected genomes by using keywords and a variety of filters, like "Gene Product name", "Locus Tag", "IMG", or "GenBank ID". In particular genes can be retrieved through a Blast search [3, 4] or performing a phylogenetic profiling.

- **Blast** (Basic Local Alignment Search Tool) functions (blastp, blastx, tblastn, blastn) allow to find matches of the selected gene sequence (with a "copy and paste" simple operation in the text box) in one or more genomes choosing the favorite e-value cutoff.
- **Phylogenic Profiler** gives to the users the possibility to analyze the phylogenetic position and the presence of homologs of a single gene or of an operon.

Single genes can also be examined with a specific function, the *Gene details* page, that includes gene, protein and pathway information, and functional predictions.

To manage every function or activity in IMG that involves **more than one gene or genome**, the user can add the genomes to the cart. For example, the Gene List (created with the addition to cart) allows the user to maintain a list of all the genes resulted from IMG analysis. After the generation of the directory (the cart), the user can upload one or more genes in the list, or export some of them in FASTA format or their information in tab Excel format.

2.3 Find Function

Functional gene study and comparisons in IMG can be performed with the button "find function". Genes can be selected using **Search item** and **Pathways** or direct links also to external browsers for functional assignment as COG, KOG, Pfam, TIGRfam, KEGG, IMG network, enzyme, phenotype, and protein family comparison.

- **Functional item** and **pathway** investigations permit to find functions in selected genomes using keywords and definite filters. In this way it is possible to restrict the search to one or few genomes that contain one gene function or metabolic pathway according to the selected functional classification.

For each available bioinformatics tool, three operations are available: *Browser*, *List*, and *List with Stats*.

Between all the available browser analysis, **IMG Networks** is also placed. Through this tool it is possible to accede to Browser

and List areas. Here, one of the most interesting functions is *IMG Pathways*, where every pathway detailed in IMG is listed. Choosing "pathway ID" button, the user can display the detail page for each one, enzymatic reactions related to that pathway, and the genomes that have at least one gene associated with the pathway (with the corresponding phylogenetic distribution).

3 Comparing Genomes on IMG

Other than being a system storage of microbial genomes, IMG is also a platform to perform **comparative analysis** of genomes. It is provided by a variety of tools that allow to compare genomes in terms of gene content, sequence conservations, clustering, synteny analysis, and distance tree. The access to the comparative analysis functions is possible from the menu options. Below some of the genome comparison tools available are presented.

- **Genome statistics** includes *summary* and *general* statistics: The *summary* comprises a variety of DNA characteristics for the selected genome, such as GC content, number of protein-coding genes, and various functional annotations, and can be summarized and split up according to COG and KEGG categories; the user can select the COG or KEGG classification links listed in the summary table, and in this way, display all the itemized data according to the selected tool. Instead, the *general* shows all the statistics for all the genomes in IMG.
- With the function **Synteny Viewers**, it is possible to visualize the DNA conservation (specifically, gene loci co-localization in different organisms) through three comparative analysis tools: *VISTA*, *Dotplot*, and *Artemis ACT*. *VISTA* is preferably used to compare sequence alignments of compared genomes to explore and study the conservation sequences. In IMG platform, a variety of pre-alignments are available for use; selecting one of the possible choices, the user can display the data alignment. *Dotplot* can generate diagrams to prospect the similarity between two or more genomes. Finally, *Artemis ACT* is used for pairwise genome DNA sequence comparisons.
- **Abundance profiles** tool permits to compare genomes in terms of abundance of protein functions and families (according to COG, Pfams, and TIGRfams). In the *Overview*, the user can view abundance for all functions of selected genomes, and can select the output would; *heat map* shows the proteins/families abundant with different colors: the red one is the most abundant. *Matrix* displays the output in tabular format. In the Search section, it is possible to research one function based on its abundance in different genomes.

- **Distance** and **Radial Tree** permit to select a minimum of three or five genomes in IMG platform ad visualize the phylogenetic tree that correlates them.
- **Genome Clustering** permits the user to clusterize genomes based on similar function profiles. During the analysis it is possible to choose the clustering method, besides genome status and upload selected sequences; the types of cluster are sorted by function (COG, Pfam, KO, TIGRfam) and by taxonomy (class, family, genus); instead the cluster methods are based on hierarchical clustering, correlation clustering, and analysis of the principal components. Most used in the analysis are *hierarchical clustering*, that shows by a tree the phylogenetic distance between compared genomes, and *correlation clustering*, that gives the possibility to display by matrix the correlation coefficient.

In the **Analysis Cart** of Genes, Functions, Genomes, and Scaffolds (incomplete genomes), the user can find all the items that he or she selected during IMG analysis. The genomes that the user want to analyze and compare can be also uploaded and, if necessary, exported and saved on personal platform.

MyIMG allows users to set preferences for platform use, and to upload and manage their genomes.

4 An Example of Pangenome Analysis with IMG

As an example of pangenome study with IMG we provide an example of comparison among genomes of the nitrogen fixing symbiotic rhizobia of genus *Ensifer*.

Browser Requirements: Java should be installed on your local OS and Java applets should be enabled.

- After accessing the IMG genome website (https://img.jgi.doe.gov) go to "Find Genome"—"Genome Search" menu.
- Type the genus name you are searching for. In this case type "Ensifer", by using as filter "Genome Name". Click the "Go" button.
- A list containing all genomes which contain the word "Ensifer" is now displayed. For each genome the taxonomic domain, the Status of the genome, the Genome Name, the Proposal Name, the Sequencing Center, the Genome size (in bp), and gene counts are reported. This view can also be customized by selecting additional search field.
- Select all genomes in the page and then click on "Add Selected to Genome cart" button.
- Go in the "Compare genome" menu and scroll down to select the different menu options.

- Select "Distance Tree" from "Compare genome" menu. The list of genomes will appear to select the total number ("Select All") or a subset of genomes (at least three).
- Click on "Select All" and then on "Go" button. A distance tree is now displayed. The tree menu allows to change fonts and the graphics of the tree. Moreover, the tree can be saved as pdf file ("Tools" - "Save as pdf") or files in phyloXML, Newick, NHX, and Nexus formats can be displayed and then copied and saved in a separate file, allowing to redraw the dendrogram with other software, as Mega [5], or TreeView (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html).
- It is possible also to proceed with a comparative analysis of functions. Firstly, you have to select one or more functions. Go to the "Find functions" menu and select one of the options, as for instance "Function Search". Here you can search for functions based on function name or on different codes (as COG, Pfam, KEGG definitions, Enzyme codes, MetaCyc). Select "Function Profile" from the "Compare Genomes" menu.
- Type "nitrite reductase" as gene product name. A list of the different nitrite reductases, their number (gene count), and the number of genomes containing each nitrite reductase is displayed. Select the copper-containing nitrite reductase. By clicking on the number of genome it is possible to visualize the genomes containing the selected gene. Cu-containing nitrite reductase is known to be part of the dispensable genome fraction of *Ensifer* sp. and confer tolerance to sodium nitrite [6–8].
- Another possibility to compare functions it to proceed with an overview of all functions. Go on the "Compare Genomes" menu and select the "Abundance Profiles" - "Overview (all functions)" option. Here you can proceed with the drawing of a heat map showing the abundance (absolute or normalized to genome size) of all functions. Functions can be chosen as COG, Enzyme, KO, Pfam, and TIGfam. For instance select "COG" and then select the genomes to be compared by searching in the list or browsing the phylogenetic tree. For instance browse the phylogenetic tree on "*Proteobacteria*", then *Alphaproteobacteria*, then *Rhizobiales*, then *Rhizobiaceae*, and finally "*Ensifer*". Select "*Ensifer*" and click the "Go" button. A heat map will be displayed with color indicating the abundance (red, high abundance; blue, low abundance) for each COG. Here functions present in all genomes (core genome) with respect to dispensable or differentially occurring functions can be identified.

5 Conclusions

This chapter has shown some key functions of IMG platform, which can be applied by also nonexperienced bioinformaticians to analyze genome data and perform several basic and advanced analyses on comparative bacterial genomics. For further information and utilities please consult the related publications [1, 9] and the manual ("Using IMG" menu on IMG webpage).

References

1. Grigoriev IV, Nordberg H, Shabalov I, Aerts A, Cantor M, Goodstein D, Kuo A, Minovitsky S, Nikitin R, Ohm RA, Otillar R, Poliakov A, Ratnere I, Riley R, Smirnova T, Rokhsar D, Dubchak I (2011) The Genome Portal of the Department of Energy Joint Genome Institute. Nucleic Acids Res 40:D26–D32
2. Markowitz VM, Chen IMA, Palaniappan K, Chu K, Szeto E, Pillay M, Ratner A, Huang J, Woyke T, Huntemann M, Anderson I, Billis K, Varghese N, Mavromatis K, Pati A, Ivanova NN, Kyrpides NC (2014) IMG 4 version of the integrated microbial genomes comparative analysis system. Nucleic Acids Res 42(D1): D560–D567
3. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402
4. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden T (2008) NCBI BLAST web site NCBI BLAST: a better web interface. Nucleic Acids Res 36:W5–W9
5. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731–2739
6. Galardini M, Mengoni A, Brilli M, Pini F, Fioravanti A, Lucas S, Lapidus A, Cheng J-F, Goodwin L, Pitluck S, Land M, Hauser L, Woyke 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
7. Biondi EG, Tatti E, Comparini D, Giuntini E, Mocali S, Giovannetti L, Bazzicalupo M, Mengoni A, Viti C (2009) Metabolic capacity of *Sinorhizobium* (*Ensifer*) *meliloti* strains as determined by phenotype microarray analysis. Appl Environ Microbiol 75:5396–5404
8. Galardini M, Mengoni A, Biondi EG, Semeraro R, Florio A, Bazzicalupo M, Benedetti A, Mocali S (2013) DuctApe: a suite for the analysis and correlation of genomic and OmniLog™ phenotype microarray data. Genomics. doi:10.1016/j.ygeno.2013.11.005
9. Markowitz VM, Mavromatis K, Ivanova NN, Chen IMA, Chu K, Kyrpides NC (2009) IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics 25:2271–2278

List of Publications

1. **Checcucci, A**., & Mengoni, A. (2015). The integrated microbial genome resource of analysis. In Mengoni, A., Galardini, M., Fondi, M. (eds.) *Bacterial Pangenomics: Methods and Protocols*. Methods in Molecular Biology Series, Volume 1211. Humana Press, Springer. pp 289-295.
2. **Checcucci, A**., Azzarello, E., Bazzicalupo, M., Galardini, M., Lagomarsino, A., Mancuso, S., ... & Zanardo, M. (2016). Mixed nodule infection in *Sinorhizobium meliloti-Medicago sativa* symbiosis suggest the presence of cheating behavior. *Frontiers in Plant Science*, 7, 835.
3. diCenzo G., **Checcucci, A**., Bazzicalupo, M., Mengoni, A., Viti, C., Dziewit, L., Finan, T. M., ... & Fondi, M. (2016). Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*. *Nature communications*, 7, Article number 1221
4. **Checcucci A**., Maida I., Bacci G., Ninno C., Bilia A.R., Biffi S., Firenzuoli F., Flamini G., Fani R., Mengoni A. (2016) Is the plant-associated microbiota of *Thymus* spp. adapted to the plant essential oil? *Research in Microbiology*. In Press.
5. **Checcucci A.,** Azzarello E., Bazzicalupo M., De Carlo A., Emiliani G., Mancuso S., Spini G., Viti C., Mengoni A. (2017) Role and regulation of ACC deaminase gene in *Sinorhizobium meliloti*: is it a symbiotic, rhizospheric or endophytic gene? *Frontiers in Genetics*, 8:6.

6. **Checcucci A**., Bazzicalupo M., Mengoni A. (2017). Exploiting nitrogen-fixing rhizobial symbionts genetic resources for improving phytoremediation of contaminated soils. In Anjum, N.A., Gill, S.S. Tuteja, N. (eds.) *Enhancing Cleanup of Environmental Pollutants: Biological Approaches*. Springer Science+Business Media (New York). In Press