Experimental systems to explore the competitiveness of root-associated bacteria and its relevance for community establishment

I n a u g u r a l - D i s s e r t a t i o n

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Mohamed Amine Hassani

aus Blida (Algerien)

Köln, Mai 2017

Die vorliegende Arbeit wurde am Max-Planck-Institut für Pflanzenzüchtungsforschung in Köln in der Abteilung für Planze-Mikroben Interaktionen (Direktor: Prof. Dr. Paul Schulze-Lefert) angefertigt.

Berichterstatter:	**Prof. Dr. Paul Schulze-Lefert**
	Prof. Dr. Alga Zuccaro
	Dr. Sebastian Fraune
Prüfungsvorsitender:	**Prof. Dr. Gunther Döhlemann**
Tag der mündlichen Prüfung:	**07.07.2017**

„No bacterium is an island“

Table of contents

Abstract

Bacteria colonize diverse ecosystems, including the plant roots. Plant root-associated bacteria derive mainly from the surrounding soil and they are important for the host growth and health. Soil properties and root-derived organic molecules are two factors that shape the plant root-associated microbiota. However, bacteria live in multispecies associations where competitive interactions do not only alter their growth, but also the community diversity, structure and stability. In this study, I assessed the role of bacteria-bacteria interactions in the establishment of bacterial communities using different experimental systems. By combining genomic, metabolomic and phenotypic studies, I explored the competitive potential of 198 bacterial isolates that are mainly derived from the roots of *Arabidopsis thaliana*. Comparative genomes analysis revealed that the bacteria harbor diverse biosynthetic gene clusters that encode enzymatic pathways for the biosynthesis of diverse specialized metabolites, including antimicrobials. The metabolomic study revealed that several bacterial strains secrete genome-predicted antimicrobials. Moreover, the screen for mutual inhibitions revealed that 66% of the isolates engage in at-distance antagonistic interactions. The screen for mutual inhibitions did not only reveal interesting phylogenetic and ecological patterns in inter-bacterial competition, but also provided a ground for defining highly competitive or highly sensitive community members. By defining two groups of bacteria with contrasting competitive potential, I tested the role of bacteria-bacteria interactions in the establishment of microbial communities in liquid microcosms and *in planta.* The perturbation experiments revealed that highly competitive strains are important for the maintenance of the community diversity and structure in liquid microcosms, whereas the *A. thaliana* root-associated microbiota were resilient to the applied perturbations. This study indicates that inter-bacterial interactions are important for the community diversity and stability in a niche-dependent manner.

Zusammenfassung

Bakterien kolonisieren verschiedene Ökosysteme, wozu auch die Pflanzenwurzeln zählen. Pflanzenwurzeln-assoziierte Bakterien sind wichtig für das Pflanzenwachstum und die Pflanzengesundheit und stammen meistens von der umliegenden Erde. Biotische Faktoren aus der Erde und aus dem Wurzelsystem stammende organische Moleküle sind zwei bekannte Faktoren, die die wurzelassoziierten Mikrobiota in Pflanzen prägen. Bakterien leben in Verbänden mit anderen Bakterien zusammen, in denen konkurrenzfähige Interaktionen nicht nur ihr Wachstum beeinflussen, sondern auch die Diversität der Gemeinschaft, die Struktur und Stabilität. In dieser Arbeit wird die Funktion von Bakterien-Bakterien Interaktionen bei der Gründung von bakteriellen Gemeinschaften in verschiedenen Ökosystemen untersucht. Durch die Kombination von genomischen, metabolomischen und phänotypischen Untersuchungen untersuchten wir das wettbewerbsfähige Potential von 198 Bakterien-Isolaten, die hauptsächlich aus den Wurzeln von Arabidopsis thaliana stammen. Vergleichende Genomanalysen zeigten, dass die Bakterien unterschiedliche biosynthetische Gencluster beherbergen, die enzymatische Signalwege für die Biosynthese verschiedenener spezifischer Metaboliten, einschließlich antimikrobieller Substanzen, kodieren. Die Metabolomanalyse zeigte, dass Bakterien antimikrobielle Substanzen sekretieren, die im Genom vorhergesagt sind. Zudem zeigte der Screen für gegenseitige Inhibitionen, dass 66% der Isolate an „at-distance“ antagonistischer Aktivität beteiligt sind. Der Screen für wechselseitige Inhibitionen zeigte nicht nur interessante phylogenetische und ökologische Muster, sondern lieferte auch den Grundstein für die Auswahl von stark wettbewerbsfähigen und stark suszeptiblen Gemeinschaftsmitgliedern. Durch die Definierung zweier Gruppen von Bakterien mit gegensätzlichem wettbewerbsfähigem Potential, testeten wir die Rolle von Bakterien-Bakterien Interaktionen bei der Gründung von mikrobiellen Gemeinschaften in flüssigem Microkosmos und in planta. Perturbationsexperimente zeigten, dass stark wettbewerbsfähige Bakterien wichtig sind für den Erhalt der Diversität und Struktur der bakteriellen Gemeinschaft im flüssigen Mikrokosmos, während wurzelassoziierte Mikrobiota in Arabidopsis belastbar gegenüber Störungen waren. Diese Arbeit liefert Hinweise darauf, dass interbakterielle Interaktionen in einer Ökosystem-abhängigen Weise wichtig für die Errichtung der bakteriellen Gemeinschaft sind.

Cooperative and competitive interactions in the microbial world and their role in altering the community diversity and stability

I.A. General Introduction

Land plants appeared approximatively 430 million years ago (Wellmann *et al.,* 2003), whereas microbes appeared much earlier (~4.2 billion years ago, Dodd *et al.,* 2017). Since then, both macro- and micro-organisms have tremendously diversified and continued to evolve in close interactions. The study of interactions between plants and microbes has propelled the emergence of a field at the crossroad of microbiology and plant physiology and revealed that plant-microbe(s) interactions span from beneficial to deleterious interactions for both organisms. Indeed, microbes colonize both below- and above-ground organs of plants where nutrients and space are available (Lambers *et al.,* 2009). Collectively, plant-colonizing microbes are referred as the plant microbiota, and is mainly dominated by bacteria, fungi, and oomycetes (Müller, 2016). Microbial communities associated with plants are generally considered to be beneficial, however it is not excluded that some microbiota members could become opportunistic pathogens (Brown *et al.,* 2012). Therefore, plant-associated microbial communities are important determinants of host health and growth in both beneficial and detrimental directions. In a beneficial manner, the plant microbiota can alter host nutrient status by providing or increasing nutrient availability (Hacquard *et al.,* 2015). Numerous studies on the molecular mechanisms of interactions between nitrogen-fixing *Rhizobium* species (Sulieman and Tran, 2014), or arbuscular mycorrhizal fungi and plants (Smith and Smith, 2011), underpin these types of beneficial interactions. Otherwise, microbes are also known to elevate plant tolerance to abiotic or biotic stresses. While certain plant-associated microbes help the host to cope with drought (Vílchez *et al.,* 2016), high salinity (Dodd and Pérez-Alfocea, 2012), or heavy metal-contaminated soil (Yang *et al.,* 2009), other microbes provide a protection against intruder phytopathogens by colonizing available space or competing against these pathogens through available resources or by a direct antagonism (Berendsen *et al.,* 2012). These few examples do not constitute an exhaustive list of studies that stress the importance of plant-microbe(s) interactions, but sufficiently highlight the microbial ability to profoundly alter not only the health and growth of their associated hosts but also ecosystem functioning.

Studying the species composition of host-associated microbial communities is crucial for understanding their functioning. The advent of high-throughput sequencing has rendered studying microbial communities associated to a host or an environment less laborious and more widespread. While ribo-sequencing offered opportunities to study the microbial composition of diverse ecosystems (inert or living) and to correlate their functioning with the community biodiversity (Delgado-Baquerizo *et al.,* 2016), metagenomic studies were efficient in inferring ecosystem functioning through the power of cataloging the gene content of microbial communities (Bulgarelli *et al.,* 2015). Unequivocally, the study of species composition or the metagenomes of ecosystem microbial communities is important.

However, the mechanisms that govern the assembly of microbial communities associated with the host are still not well understood. More precisely, the role of inter-microbial interactions in the assembly and stability of the plant microbiota is far less understood. The lack of in-depth studies that explore the mechanisms involved in microbiota assemblage is due to the complexity of microbial communities associated with the host and to the unpredictability of the surrounding environment. In order to deconvolute host-associated microbiota complexity, it is useful to work with simplified synthetic microbial communities that are representative of the host microbiota, and to employ gnotobiotic experimental systems that mimic natural habitats but under controlled growth conditions. In the recent years, several studies employed synthetic microbial communities isolated from plants and empirically test working hypotheses regarding the assembly of microbial communities on or in the host under controlled laboratory conditions (Bodenhausen *et al.*, 2014, Bai *et al.*, 2015, Lebeis *et al.*, 2015, Niu *et al.*, 2017). Using microbial synthetic communities that resemble natural communities is an alternative and reductionist approach that allows one to gradually increase species complexity, empirically test hypotheses, and to perform perturbation experiments under controlled abiotic parameters.

I.B. The plant root bacterial microbiota

Bacteria are important members of the plant microbiota, which have been shown to alter significantly plant health and growth (Turner *et al.*, 2013). While land plants use soil as a matrix to support their growth and uptake minerals, a subset of soil derived bacteria engage in interactions with the plants *via* their roots. Although very limited in nutrients, soil as an ecosystem contains tremendous bacterial diversity and only a small fraction (~0.1-1%) of its bacteria have been grown *in vitro* (Amann *et al.*, 1995). Soil physical and chemical properties are key for defining soil resident biota (Kim *et al.*, 2014). In contrast to soil, the rhizosphere (the vicinity of the plant root) and the rhizoplane (plant root surface) are known to be densely colonized by bacteria from predominantly three phyla, *Proteobacteria*, *Bacteroidetes* and *Actinobacteria* (with *Firmicutes* at lower percentage), as revealed by culture-independent profiling of several plant species (Dombrowski *et al.*, 2017, Edwards *et al.*, 2015, Peiffer *et al.*, 2013, Lundberg *et al.*, 2012, Bulgarelli *et al.*, 2012). Interestingly, the establishment of the plant root microbiota occurs in the early days after seed germination and remains relatively stable afterwards (Edwards *et al.*, 2015). The establishment of plant-associated microbiota is suggested to be initiated by root exudation. Therefore, both soil edaphic characteristics and root-derived organic molecules influence the establishment of plant-associated bacterial microbiota as referred to by a “two-step” selection model (Bulgarelli *et al.*, 2012). Nonetheless, it is also important to acknowledge that dispersion and speciation are two contributing additional factors that influence microbiota establishment (Herrera Paredes and Lebeis, 2016, Nemergut *et al.*, 2013).

Since plants release up to 21% of photosynthetically fixed carbon through root exudation (Badri and Vivanco, 2009), it is not surprising that a large proportion of copiotroph bacterial species colonize plant rhizosphere or rhizoplane in order to escape from competition in the soil. In contrast, slow-growing

oligotrophic bacteria such as *Acidobacteria* grow better in carbon-poor environments and are therefore often out-competed in the roots by copiotroph bacteria (Fierer *et al.,* 2007). Remarkably, isolation efforts from *Arabidopsis* roots conducted by Bai and colleagues have realized up to ~66% recovery of bacterial species that have been reported as root-associated bacteria based on amplicon sequencing (Bai *et al.,* 2015). According to these data, it is plausible to assume that most root-associated bacteria may be copiotroph species that can be isolated under laboratory conditions. Undoubtedly, these data also indicate that plants are a "hotspot" for nutrients where bacterial species meet and mix. Since the roots are rich in nutrients compared to bulk soil and densely colonized by bacteria, it is conceivable that a plethora of ecological interactions between bacteria occur in the plant rhizoplane or rhizosphere. As bacteria-bacteria interactions vary on the spectrum of cooperative to competitive interactions, these interactions likely influence the reproduction and the survival of interacting species (Hibbing *et al.,* 2010). Furthermore, bacterial interactions can influence community structure and diversity (Stubbendieck *et al.,* 2016) which subsequently could affect host fitness (Fraune *et al.,* 2015). In the following, several mechanisms of cooperative and competitive mechanisms are briefly reviewed in the context of plant-associated bacteria.

I.C. Cooperative interactions for the common good

Cooperative interactions are beneficial reciprocal interactions that involve closely or distantly related species (Freilich *et al.,* 2011). Syntrophy is a cooperative mechanism that benefits both involved cells and occurs between metabolically interdependent bacteria (Morris *et al.,* 2013). Through metabolic interdependency, bacteria can extend their fundamental niche to face nutrient-poor environments (Zelezniak *et al.,* 2015), break down recalcitrant compounds (Westerholm *et al.*,2011), remove toxic metabolites, exchange electrons, or exchange organic, sulfurous or nitrogenous compounds (Schink, 2002). Microbial interdependencies are one consequence of reductive evolution that is explained by gene or function loss in the interacting partner cells (Morris *et al.,* 2012). Although common in the microbial world, syntrophic interactions among plant-associated bacteria have not been yet revealed.

Bacteria also cooperate through secretion of "public goods" molecules that benefit all community members in the environs, regardless of whether all community members are contributing to the community functioning or not (Griffin *et al.,* 2004). Secretion of "public goods" can be critical for both the microbe and the host. For example, bacteria are known to secrete an extracellular polymeric substance (EPS) matrix, which is a "public good", in order to build biofilms (Stoodley *et al.,* 2002). A bacterial biofilm is an assemblage of microbial communities that adhere to a surface and grow embedded in an extracellular matrix (Donlan, 2002). Biofilm is a micro-architectural construction that requires clonal or multispecies cooperation in order to be achieved and maintained (López *et al.,* 2010). The ability to form a biofilm could be advantageous for plant associated-bacteria since it provides protection against microbial competitors or secreted antimicrobial molecules (Van Acker *et al.,* 2014). Importantly,

bacteria inside a biofilm can perform enzymatic processes that require high cellular density (Nadell *et al.*, 2008) or acquire new functions *via* horizontal gene transfer (Branda *et al.*, 2005). Several studies have demonstrated that plant associated-bacteria can form biofilms on plant tissue (Danhorn and Fuqua, 2007, Bogino *et al.*, 2013). Biofilm-inhabiting bacteria can be beneficial or deleterious for plant fitness (Danhorn and Fuqua, 2007). For instance, *Sinorhizobium meliloti*, a beneficial microbe, requires the secretion of biofilm-forming exopolymers for effective symbiosis with legume plants (Fujishige *et al.*, 2006). In contrast, the pathogen *Pseudomonas aeruginosa* has been reported to form a biofilm on sweet basil roots to protect itself from antimicrobial compounds secreted by the plant (Walker *et al.*, 2004).

To build microbial biofilm, bacteria must coordinate and/or synchronize their secretions. Coordination of group behavior is mediated through molecular communication (Greenberg, 2003). Quorum sensing is a process that allow bacteria to monitor their populations and coordinate processes through the secretion of diffusible molecules (Davies *et al.*, 1999, Fuqua *et al.*, 2001). Whereas both Gram-positive and Gram-negative bacteria practice quorum sensing, the employed interspecies communication mechanisms in these two groups are fundamentally different (Federle and Bassler, 2003). Coordination of group behavior allows bacteria to perform functions that are relevant only at high cellular density. For instance, the phytopathogen *Erwinia carotovora* secretes cell wall-degrading exo-enzymes upon reaching high cellular density. The large scale secretion is required for bacterial pathogenicity and allows the phytopathogen to overwhelm plant defenses (Pirhonen *et al.*, 1993). Coordination of cellular group behavior is also known to help bacteria to overcome environmental stresses such as those caused by iron deficiency. Iron is a limiting nutrient in soil and is essential for many cellular process (Stintzi *et al.*, 1998). Bacteria secrete chelating agents to scavenge and make iron available to the cell (Neilands, 1995) and several studies have elegantly demonstrated how plants can use microbial siderophores to alleviate environmental iron-deficiencies (Crowley *et al.*, 1992, Fernandez *et al.*, 2005, Radzki *et al.*, 2013, Trapet *et al.*, 2016). Importantly, secreted siderophores are also utilized by community members that may partially or not participating to the siderophore production (Griffin *et al.*, 2004). Secretion of public goods fosters social interactions among community members, however it also favors the emergence of "social cheater" bacteria that gain in fitness by escaping the shared fair cost due to the production of the metabolites (West *et al.*, 2007). Consequently, "social cheater" bacteria can increase in frequency and out-compete cooperating species, thereby jeopardizing the community and its functioning (Velicier *et al.*, 2000, Griffin *et al.*, 2004, Sandroz *et al.*, 2007).

Even though cooperative interactions are common among microorganisms, the evolution and maintenance of such interactions remains puzzling for evolutionary biologists. Recently, the black queen hypothesis emerged as model explaining observed interdependencies between microbes (Morris *et al.*, 2012). Several theoretical and empirical studies helped us understanding the evolution and maintenance of cooperative interactions, though often these studies consider only one trait. Key factors have been proposed as important in the evolution of cooperative interactions and their maintain: 1- cost and benefits from the interactions, 2- relatedness of the interacting members, and 3- fidelity of interacting

cells (West *et al.*,2007, Foster and Wenseleers, 2005, Ross-Gillespie *et al.*,2015). Although understanding the maintenance of cooperation is interesting from the evolutionary point of view, little attention has been given to the role of cooperative interactions in maintaining community stability. Only recently, a theoretical-based study has suggested that predominance of cooperative interactions tend to destabilize microbial communities (Coyte *et al.,* 2015). In contrast, competitive interactions are proposed to promote community diversity and stability (Coyte *et al.*, 2015, Stubbendieck *et al.*, 2016). Interestingly, competitive interactions are recognized as prevalent (Pérez-Gutiérrez *et al.,* 2013) and constitute a major component of inter-bacterial interactions (Foster and Bell, 2012, Oliveira *et al.*, 2014).

I.D. Competitive interactions for the greater good of the community diversity and stability

Ecological competition refers to the biological interactions between same or different species that aim to negatively alter survival or reproduction of opponent cells. Competitive interactions are important ecological factors that affect bacterial community diversity (Czárán *et al.*, 2002), spatial structure (Kim *et al.,* 2008), stability (Keslic *et al.,* 2015), and ultimately functioning (Wei *et al.,* 2015). Indeed, theoretical and empirical studies have showed that competition can promote species' evolutionary diversification through resource use diversity and/or promoting spatial structure (Kerr *et al.*, 2002, Day and Young, 2004, Grether *et al.*, 2009, Svanbäck and Bolnick, 2006). Therefore, the combination of both biotic and abiotic factors shapes bacterial community structures and their evolutionary trajectories (Tenaillon *et al.,* 2012, Khare and Tavazoie, 2015). A constant battle for resources under fluctuating environments has likely contributed to the emergence of diverse competitive mechanisms that bacteria employ to protect territory or to conquer new niches. Two main categories of competitive interactions are distinguished within the bacterial world: exploitative competition and interference competition (Hibbing *et al.*, 2010); these are briefly discussed below.

Exploitative competitive interactions are indirect competitive mechanisms mediated by rapid and efficient utilization of limiting resources. Bacteria use several sophisticated exploitative mechanisms to compete against their neighbors. A good example of exploitative competition is illustrated by rapid and efficient iron sequestration *via* secretion of bacterial siderophores (Wandersman and Delepelaire, 2004). Under iron-limiting condition, bacteria secrete siderophores that drastically reduce iron availability in the environs, impeding subsequently the growth of other microbes (Chu *et al.,* 2010). Siderophores are molecules with low molecular weight secreted by bacteria for iron solubilization, transport, and storage (Hider and Kong, 2010). Well-documented examples of soil bacterial siderophores include the secretion of pyoverdines by *Pseudomonas* (Trapet *et al.,* 2016), ornibactin by *Burkholderia* (Deng *et al.,* 2017), and bacillibactin by *Bacillus* (Chowdhury *et al.,* 2015). Nutrient sequestration is recognized as an important trait in biocontrol bacteria to out-compete pathogens (Whipps, 2001, Friesen *et al.,* 2011) and has been linked to the suppression of diseases caused by fungal or bacterial pathogens (Suárez-Estrella

et al., 2013). Furthermore, it has been recently shown that resource competition is an important factor linking bacterial community composition and pathogen invasion in the rhizosphere of tomato plants (Wei *et al.*, 2015). These results not only underline the role of resource competition for microbial interactions but also indicate their relevance for plant health. Interestingly, exploitative competitive interactions are proposed to limit and influence bacterial population size and are predicted to lead to interference competitive interactions (Little *et al.*, 2008, Cornforth and Foster, 2013, Holdridge *et al.*, 2016).

Interference competitive interactions refer to competitive mechanisms that involve direct harm of opponent bacteria. These interactions are intended to suppress the growth of opponent cells *via* contact-dependent and/or contact-independent mechanisms (Hibbing *et al.*, 2010). Contact-dependent interference is an inhibitory mechanism that requires a direct contact between competing cells and is mediated by diverse molecular systems, including the type V secretion system (Aoki *et al.*,2005, reviewed in Ruhe *et al.*,2013), the type VI secretion system, (Basler *et al.*,2014), and the *Rhs* system (Poole *et al.*,2011). Plant-associated bacteria have been reported to engage in direct antagonistic interactions mediated by contact-dependent killing mechanisms. For example, the plant pathogen *Agrobacterium tumefaciens* uses a puncturing type VI secretion system to deliver DNase effectors upon contact with a bacterial competitor *in vitro* and *in planta*. Remarkably, this contact-dependent antagonism provides a fitness advantage for the bacterium only *in planta*, underlining its specific importance for niche colonization in a natural habitat (Ma *et al.*, 2014).

In contrast to contact-dependent inhibition, bacteria also employ at-a-distance killing mechanisms mediated by diffusible or volatile metabolites (Tyc *et al.*, 2017). Former metabolites are polar compounds that diffuse through liquids and known for their potent inhibitory activity. Several classes of diffusible antimicrobials mediate interference competition. For instance, bacteria are known to secrete ribosomally-synthesized peptides known as bacteriocins. Bacteriocins exhibit similar antimicrobial activity against closely related species as against distantly related species (Abrudan *et al.*, 2012). Plant-associated bacteria that produce bacteriocins can confer to the host protection against pathogens (Subramanian and Smith, 2015). For instance, *Bacillus subtilis* inhibits the growth of the pathogen *Agrobacterium tumefaciens* through the secretion of a bacteriocin (Hammami *et al.*, 2009). In contrast to bacteriocins, bacteria also secrete nonribosomal peptides (nrps) or polyketides (pks) that have broad biological activity and include some notoriously potent antibiotics like teixobactin (Losee *et al.*, 2015) or erythromycin (Shen, 2003). The latter metabolites are gaseous organic compounds that diffuse easily through air and can be organic (Schulz and Dickschat, 2007). Volatile organic compounds (VOCs) include diverse classes of molecules that are ideal for long-distance inter-bacterial interactions (Tyc *et al.*, 2015). Several classes of VOCs are produced by soil or rhizospheric bacteria, like terpenes, pyrazines or indole (Tyc *et al.*, 2017). The study of volatile compounds as mediators of inter-bacterial antagonism is nascent. A recent study has shown that *Collimonas pratensis* produces a blend of sesquiterpenes that have antimicrobial activity against *Staphylococcus aureus* and *Escherichia coli*

(Song *et al.,* 2015). However, it is still unknown whether such volatile compounds are used by plant-associated bacteria to inhibit plant pathogens.

Antagonistic interactions have an important role in bacterial competitiveness. Secretion of antibiotic peptides have been shown to provide a competitive advantage for *Rhizobium etli* strains inside root nodules (Robleto *et al.,* 1998). As it is becoming clear that the production of antimicrobials by plant-associated bacteria benefits plant health, plant pathogenic bacteria also enhance their own success by the secretion of antimicrobial molecules. The tuber pathogen, *Clostridium puniceum*, secretes antimicrobial polyketides called clostrubins in order to compete against other bacteria and further invade aerobic micro-habitats (Shabuer *et al.,* 2015). Undoubtedly, antibiosis warfare is an important mediator of intra- and inter-specific competitive interactions, but its frequency among microbiota community members has been rarely explored. Only a handful of studies, so far, have examined antagonistic interactions between an ecosystem-associated bacteria. For instance, Rypien and colleagues tested antagonistic interactions among the bacterial strains isolated from the scleractinian coral, *Montastrea annularis*, at two different temperatures (Rypien *et al.,* 2010). The authors demonstrated that antagonistic interactions are prevalent among coral-associated bacteria and, further, indicated their importance for community diversity and spatial heterogeneity (Rypien *et al.,* 2010). More recently, the study of antagonistic interactions between bacterial isolates from the rhizosphere, roots, and phyllosphere of the medicinal plant, *Echinacea purpurea*, have suggested that plant-associated bacteria compete against each other through the secretion of antimicrobials (Maida *et al.,* 2015). Moreover, bacteria from the different plant compartments showed different levels of sensitivity to antagonistic activity, thereby suggesting that antagonistic interactions play an important role in shaping the structure of the plant microbiota (Maida *et al.,* 2015). However, to what extent competitive interactions between plant-associated microbiota members alter the community diversity and structure have not been yet tested *in vitro* or *in planta*.

I.E. Objectives of the research project

Competitive interactions mediated by antimicrobials are important for the host and the host-associated microbial communities. Indeed competitive interactions contribute to the host protection against intruder pathogens (Whipps, 2001). Importantly, inter-bacterial competition also promote the microbial community diversity (Czárán *et al.,* 2002), spatial structure (Narisawa *et al.,* 2008) and stability (Coyte *et al.,* 2015). However, to what extent competitive interactions contribute to shaping the host-associated microbiota is not well described. This study aim at exploring the competitive potential of several *Arabidopsis thaliana* root-associated bacterial strains in order to define highly competitive community member and test the role of these bacteria in altering the community diversity and structure in liquid microcosms and *in planta.*

In order to explore the competitive potential of microbiota community members, I combined genomic, metabolomic and phenotypic approaches. These approaches aim at; i) *in silico* identification of gene clusters for the biosynthesis of antimicrobials, ii) the analysis of the bacterial metabolites in order to identify secreted antimicrobials, iii) the screen for inter-bacterial antagonistic interactions mediated by at-distance antagonistic interactions and to define highly or highly sensitive strains.

To test the role of bacteria-bacteria interactions in altering the community diversity and structure, I performed community perturbation experiments by depleting 13 highly competitive or highly sensitive community members. The aim of these perturbation experiments is to reveal the role of highly competitive or highly sensitive bacteria in altering the community diversity and stability in liquid microcosms and *in planta*.

I.F. References

- Abrudan, M. I., Smakman, F., Grimbergen, A. J., Westhoff, S., Miller, E. L., van Wezel, G. P., et al. (2015). Socially mediated induction and suppression of antibiosis during bacterial coexistence. *Proc Natl Acad Sci U S A* 112, 11054–11059. doi:10.1073/pnas.1504076112.
- Amann, R. I., Ludwig, W., and Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59, 143–169.
- Aoki, S. K., Pamma, R., Hernday, A. D., Bickham, J. E., Braaten, B. A., and Low, D. A. (2005). Contact-dependent inhibition of growth in Escherichia coli. *Science* 309, 1245–1248. doi:10.1126/science.1115109.
- Badri, D. V., and Vivanco, J. M. (2009). Regulation and function of root exudates. *Plant Cell Environ.* 32, 666–681. doi:10.1111/j.1365-3040.2008.01926.x.
- Bai, Y., Müller, D. B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M., et al. (2015). Functional overlap of the Arabidopsis leaf and root microbiota. *Nature* 528, 364–369. doi:10.1038/nature16192.
- Basler, M., Ho, B. T., and Mekalanos, J. J. (2013). Tit-for-tat: Type VI secretion system counterattack during bacterial cell-cell interactions. *Cell* 152, 884–894. doi:10.1016/j.cell.2013.01.042.
- Berendsen, R. L., Pieterse, C. M. J., and Bakker, P. A. H. M. (2012). The rhizosphere microbiome and plant health. *Trends Plant Sci.* 17, 478–486. doi:10.1016/j.t plants.2012.04.001.
- Bodenhausen, N., Bortfeld-Miller, M., Ackermann, M., and Vorholt, J. A. (2014). A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. *PLoS Genet.* 10, e1004283. doi:10.1371/journal.pgen.1004283.
- Bogino, P. C., de las Mercedes Oliva, M., Sorroche, F. G., and Giordano, W. (2013). The Role of Bacterial Biofilms and Surface Components in Plant-Bacterial Associations. *Int. J. Mol. Sci.* 14, 15838–15859. doi:10.3390/ijms140815838.
- Branda, S. S., Vik, Å., Friedman, L., and Kolter, R. (2005). Biofilms: the matrix revisited. *Trends in Microbiology* 13, 20–26. doi:10.1016/j.tim.2004.11.006.
- Brown, S. P., Cornforth, D. M., and Mideo, N. (2012). Evolution of virulence in opportunistic pathogens: generalism, plasticity, and control. *Trends Microbiol.* 20, 336–342. doi:10.1016/j.tim.2012.04.005.
- Bulgarelli, D., Garrido-Oter, R., Münch, P. C., Weiman, A., Dröge, J., Pan, Y., et al. (2015). Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* 17, 392–403. doi:10.1016/j.chom.2015.01.011.
- Bulgarelli, D., Rott, M., Schlaeppi, K., Ver Loren van Themaat, E., Ahmadinejad, N., Assenza, F., et al. (2012). Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* 488, 91–95. doi:10.1038/nature11336.

- Bulgarelli, D., Schlaeppi, K., Spaepen, S., Ver Loren van Themaat, E., and Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol* 64, 807–838. doi:10.1146/annurev-arplant-050312-120106.
- Chowdhury, S. P., Hartmann, A., Gao, X., and Borriss, R. (2015). Biocontrol mechanism by root-associated Bacillus amyloliquefaciens FZB42 – a review. *Front Microbiol* 6. doi:10.3389/fmicb.2015.00780.
- Chu, B. C., Garcia-Herrero, A., Johanson, T. H., Krewulak, K. D., Lau, C. K., Peacock, R. S., et al. (2010). Siderophore uptake in bacteria and the battle for iron with the host; a bird's eye view. *Biometals* 23, 601–611. doi:10.1007/s10534-010-9361-x.
- Cornforth, D. M., and Foster, K. R. (2013). Competition sensing: the social side of bacterial stress responses. *Nat Rev Micro* 11, 285–293. doi:10.1038/nrmicro2977.
- Coyte, K. Z., Schluter, J., and Foster, K. R. (2015). The ecology of the microbiome: Networks, competition, and stability. *Science* 350, 663–666. doi:10.1126/science.aad2602.
- Crowley, D., Romheld, V., Marschner, H., and Szaniszlo, P. (1992). Root-Microbial Effects on Plant Iron Uptake from Siderophores and Phytosiderophores. *Plant Soil* 142, 1–7.
- Czárán, T. L., Hoekstra, R. F., and Pagie, L. (2002). Chemical warfare between microbes promotes biodiversity. *Proc Natl Acad Sci U S A* 99, 786–790. doi:10.1073/pnas.012399899.
- Danhorn, T., and Fuqua, C. (2007). Biofilm Formation by Plant-Associated Bacteria. *Annual Review of Microbiology* 61, 401–422. doi:10.1146/ annu. rev. micro. 61.080706.093316.
- Davies, D. G., Parsek, M. R., Pearson, J. P., Iglewski, B. H., Costerton, J. W., and Greenberg, E. P. (1998). The Involvement of Cell-to-Cell Signals in the Development of a Bacterial Biofilm. *Science* 280, 295–298. doi:10.1126/science.280.5361.295.
- Day, T., and Young, K. A. (2004). Competitive and Facilitative Evolutionary Diversification. *BioScience* 54, 101–109. doi:10.1641/0006-3568 (2004) 054[0101:CAFED] 2.0.CO;2.
- Delgado-Baquerizo, M., Maestre, F. T., Reich, P. B., Jeffries, T. C., Gaitan, J. J., Encinar, D., et al. (2016). Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nat Commun* 7, 10541. doi:10.1038/ncomms10541.
- Deng, P., Foxfire, A., Xu, J., Baird, S. M., Jia, J., Delgado, K. H., et al. (2017). Siderophore product ornibactin is required for the bactericidal activity of Burkholderia contaminans MS14. *Appl. Environ. Microbiol.*, AEM. 00051-17. doi:10.1128/ AEM.00051-17.
- Dodd, I. C., and Pérez-Alfocea, F. (2012). Microbial amelioration of crop salinity stress. *J. Exp. Bot.* 63, 3415–3428. doi:10.1093/jxb/ers033.
- Dodd, M. S., Papineau, D., Grenne, T., Slack, J. F., Rittner, M., Pirajno, F., et al. (2017). Evidence for early life in Earth's oldest hydrothermal vent precipitates. *Nature* 543, 60–64. doi:10.1038/nature21377.
- Dombrowski, N., Schlaeppi, K., Agler, M. T., Hacquard, S., Kemen, E., Garrido-Oter, R., et al. (2017). Root microbiota dynamics of perennial Arabis alpina are dependent on soil residence

time but independent of flowering time. *ISME J* 11, 43–55. doi:10.1038/ismej.2016.109.

- Donlan, R. M. (2002). Biofilms: microbial life on surfaces. *Emerging Infect. Dis.* 8, 881–890. doi:10.3201/eid0809.020063.
- Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N. K., Bhatnagar, S., et al. (2015). Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl. Acad. Sci. U.S.A.* 112, E911-920. doi:10.1073/ pnas.1414592112.
- Federle, M. J., and Bassler, B. L. (2003). Interspecies communication in bacteria. *J. Clin. Invest.* 112, 1291–1299. doi:10.1172/JCI20195.
- Fernandez, V., Ebert, G., and Winkelmann, G. (2005). The use of microbial siderophores for foliar iron application studies. *Plant Soil* 272, 245–252. doi:10.1007/s11104-004-5212-2.
- Fierer, N., Bradford, M. A., and Jackson, R. B. (2007). Toward an Ecological Classification of Soil Bacteria. *Ecology* 88, 1354–1364. doi:10.1890/05-1839.
- Foster, K. R., and Bell, T. (2012). Competition, Not Cooperation, Dominates Interactions among Culturable Microbial Species. *Current Biology* 22, 1845–1850. doi:10.1016/j.cub.2012.08.005.
- Foster, K. R., and Wenseleers, T. (2006). A general model for the evolution of mutualisms. *Journal of Evolutionary Biology* 19, 1283–1293. doi:10.1111/j.1420-9101.2005.01073.x.
- Fraune, S., Anton-Erxleben, F., Augustin, R., Franzenburg, S., Knop, M., Schröder, K., et al. (2015). Bacteria-bacteria interactions within the microbiota of the ancestral metazoan Hydra contribute to fungal resistance. *ISME J* 9, 1543–1556. doi:10.1038/ismej.2014.239.
- Freilich, S., Zarecki, R., Eilam, O., Segal, E. S., Henry, C. S., Kupiec, M., et al. (2011). Competitive and cooperative metabolic interactions in bacterial communities. *Nat Commun* 2, 589. doi:10.1038/ncomms1597.
- Friesen, M. L., Porter, S. S., Stark, S. C., Wettberg, E. J. von, Sachs, J. L., and Martinez-Romero, E. (2011). Microbially Mediated Plant Functional Traits. *Annual Review of Ecology, Evolution, and Systematics* 42, 23–46. doi:10.1146/annurev-ecolsys-102710-145039.
- Fujishige, N. A., Kapadia, N. N., De Hoff, P. L., and Hirsch, A. M. (2006). Investigations of Rhizobium biofilm formation. *FEMS Microbiol. Ecol.* 56, 195–206. doi:10.1111/j.1574 6941.2005.00044.x.
- Fuqua, C., Parsek, M. R., and Greenberg, E. P. (2001). Regulation of Gene Expression by Cell-to-Cell Communication: Acyl-Homoserine Lactone Quorum Sensing. *Annual Review of Genetics* 35, 439–468. doi:10.1146/annurev.genet.35.102401.090913.
- Greenberg, E. P. (2003). Bacterial communication and group behavior. *J. Clin. Invest.* 112, 1288– 1290. doi:10.1172/JCI20099.
- Grether, G. F., Losin, N., Anderson, C. N., and Okamoto, K. (2009). The role of interspecific interference competition in character displacement and the evolution of competitor recognition. *Biological Reviews* 84, 617–635. doi:10.1111/j.1469-185X.2009.00089.x.

- Griffin, A. S., West, S. A., and Buckling, A. (2004). Cooperation and competition in pathogenic bacteria. *Nature* 430, 1024–1027. doi:10.1038/nature02744.
- Hacquard, S., Garrido-Oter, R., González, A., Spaepen, S., Ackermann, G., Lebeis, S., et al. (2015). Microbiota and Host Nutrition across Plant and Animal Kingdoms. *Cell Host Microbe* 17, 603–616. doi:10.1016/j.chom.2015.04.009.
- Hammami, I., Rhouma, A., Jaouadi, B., Rebai, A., and Nesme, X. (2009). Optimization and biochemical characterization of a bacteriocin from a newly isolated Bacillus subtilis strain 14B for biocontrol of Agrobacterium spp. strains. *Lett. Appl. Microbiol.* 48, 253–260. doi:10.1111/j.1472-765X.2008.02524.x.
- Herrera Paredes, S., and Lebeis, S. L. (2016). Giving back to the community: microbial mechanisms of plant–soil interactions. *Funct Ecol* 30, 1043–1052. doi:10.1111/1365-2435.12684.
- Hibbing, M. E., Fuqua, C., Parsek, M. R., and Peterson, S. B. (2010). Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol* 8, 15–25. doi:10.1038/nrmicro2259.
- Hider, R. C., and Kong, X. (2010). Chemistry and biology of siderophores. *Nat Prod Rep* 27, 637–657. doi:10.1039/b906679a.
- Holdridge, E. M., Cuellar-Gempeler, C., and terHorst, C. P. (2016). A shift from exploitation to interference competition with increasing density affects population and community dynamics. *Ecol Evol* 6, 5333–5341. doi:10.1002/ece3.2284.
- Kelsic, E. D., Zhao, J., Vetsigian, K., and Kishony, R. (2015). Counteraction of antibiotic production and degradation stabilizes microbial communities. *Nature* 521, 516–519. doi:10.1038/nature14485.
- Kerr, B., Riley, M. A., Feldman, M. W., and Bohannan, B. J. M. (2002). Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* 418, 171–174. doi:10.1038/nature00823.
- Khare, A., and Tavazoie, S. (2015). Multifactorial Competition and Resistance in a Two-Species Bacterial System. *PLoS Genet.* 11, e1005715. doi:10.1371/journal.pgen.1005715.
- Kim, H. J., Boedicker, J. Q., Choi, J. W., and Ismagilov, R. F. (2008). Defined spatial structure stabilizes a synthetic multispecies bacterial community. *PNAS* 105, 18188–18193. doi:10.1073/pnas.0807935105.
- Kim, H. M., Jung, J. Y., Yergeau, E., Hwang, C. Y., Hinzman, L., Nam, S., et al. (2014). Bacterial community structure and soil properties of a subarctic tundra soil in Council, Alaska. *FEMS Microbiol. Ecol.* 89, 465–475. doi:10.1111/1574-6941.12362.
- Lambers, H., Mougel, C., Jaillard, B., and Hinsinger, P. (2009). Plant-microbe-soil interactions in the rhizosphere: an evolutionary perspective. *Plant Soil* 321, 83–115. doi:10.1007/s11104-009-0042-x.

- Lebeis, S. L., Paredes, S. H., Lundberg, D. S., Breakfield, N., Gehring, J., McDonald, M., et al. (2015). PLANT MICROBIOME. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* 349, 860–864. doi:10.1126/science.aaa8764.
- Ling, L. L., Schneider, T., Peoples, A. J., Spoering, A. L., Engels, I., Conlon, B. P., et al. (2015). A new antibiotic kills pathogens without detectable resistance. *Nature* 517, 455–459. doi:10.1038/nature14098.
- Little, A. E. F., Robinson, C. J., Peterson, S. B., Raffa, K. F., and Handelsman, J. (2008). Rules of Engagement: Interspecies Interactions that Regulate Microbial Communities. *Annual Review of Microbiology* 62, 375–401. doi:10.1146/ annu. rev. micro. 030608.101423.
- López, D., Vlamakis, H., and Kolter, R. (2010). Biofilms. *Cold Spring Harb Perspect Biol* 2. doi:10.1101/cshperspect.a000398.
- Lundberg, D. S., Lebeis, S. L., Paredes, S. H., Yourstone, S., Gehring, J., Malfatti, S., et al. (2012). Defining the core Arabidopsis thaliana root microbiome. *Nature* 488, 86–90. doi:10.1038/nature11237.
- Ma, L.-S., Hachani, A., Lin, J.-S., Filloux, A., and Lai, E.-M. (2014). Agrobacterium tumefaciens Deploys a Superfamily of Type VI Secretion DNase Effectors as Weapons for Interbacterial Competition In Planta. *Cell Host Microbe* 16, 94–104. doi:10.1016/j.chom.2014.06.002.
- Maida, I., Chiellini, C., Mengoni, A., Bosi, E., Firenzuoli, F., Fondi, M., et al. (2015). Antagonistic interactions between endophytic cultivable bacterial communities isolated from the medicinal plant Echinacea purpurea. *Environ Microbiol*, n/a-n/a. Doi:10.1111/1462 2920.12911.
- Morris, B. E. L., Henneberger, R., Huber, H., and Moissl-Eichinger, C. (2013). Microbial syntrophy: interaction for the common good. *FEMS Microbiology Reviews* 37, 384–406. doi:10.1111/1574- 6976.12019.
- Morris, J. J., Lenski, R. E., and Zinser, E. R. (2012). The Black Queen Hypothesis: Evolution of Dependencies through Adaptive Gene Loss. *mBio* 3, e00036-12. doi:10.1128/mBio.00036-12.
- Müller, D. B., Vogel, C., Bai, Y., and Vorholt, J. A. (2016). The Plant Microbiota: Systems-Level Insights and Perspectives. *Annu. Rev. Genet.* 50, 211–234. doi:10.1146/annurev-genet-120215-034952.
- Nadell, C. D., Xavier, J. B., and Foster, K. R. (2009). The sociobiology of biofilms. *FEMS Microbiology Reviews* 33, 206–224. doi:10.1111/j.1574-6976.2008.00150.x.
- Narisawa, N., Haruta, S., Arai, H., Ishii, M., and Igarashi, Y. (2008). Coexistence of Antibiotic-Producing and Antibiotic-Sensitive Bacteria in Biofilms Is Mediated by Resistant Bacteria. *Appl Environ Microbiol* 74, 3887–3894. doi:10.1128/AEM.02497-07.
- Neilands, J. (1995). Siderophores - Structure and Function of Microbial Iron Transport

Compounds. *J. Biol. Chem.* 270, 26723–26726.

- Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., et al. (2013). Patterns and Processes of Microbial Community Assembly. *Microbiol Mol Biol Rev* 77, 342–356. doi:10.1128/MMBR.00051-12.
- Niu, B., Paulson, J. N., Zheng, X., and Kolter, R. (2017). Simplified and representative bacterial community of maize roots. *Proc. Natl. Acad. Sci. U.S.A.* 114, E2450–E2459. doi:10.1073/pnas.1616148114.
- Oliveira, N. M., Niehus, R., and Foster, K. R. (2014). Evolutionary limits to cooperation in microbial communities. *PNAS* 111, 17941–17946. doi:10.1073/pnas.1412673111.
- Peiffer, J. A., Spor, A., Koren, O., Jin, Z., Tringe, S. G., Dangl, J. L., et al. (2013). Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6548–6553. doi:10.1073/pnas.1302837110.
- Pérez-Gutiérrez, R.-A., López-Ramírez, V., Islas, Á., Alcaraz, L. D., Hernández-González, I., Olivera, B. C. L., et al. (2013). Antagonism influences assembly of a Bacillus guild in a local community and is depicted as a food-chain network. *ISME J* 7, 487–497. doi:10.1038/ismej.2012.119.
- Pirhonen, M., Flego, D., Heikinheimo, R., and Palva, E. (1993). A Small Diffusible Signal Molecule Is Responsible for the Global Control. *Embo J.* 12, 2467–2476.
- Poole, S. J., Diner, E. J., Aoki, S. K., Braaten, B. A., Roodenbeke, C. t'Kint de, Low, D. A., et al. (2011). Identification of Functional Toxin/Immunity Genes Linked to Contact-Dependent Growth Inhibition (CDI) and Rearrangement Hotspot (Rhs) Systems. *PLOS Genetics* 7, e1002217. doi:10.1371/journal.pgen.1002217.
- Radzki, W., Gutierrez Mañero, F. J., Algar, E., Lucas García, J. A., García-Villaraco, A., and Ramos Solano, B. (2013). Bacterial siderophores efficiently provide iron to iron-starved tomato plants in hydroponics culture. *Antonie Van Leeuwenhoek* 104, 321–330. doi:10.1007/s10482-013-9954-9.
- Robleto, E. A., Kmiecik, K., Oplinger, E. S., Nienhuis, J., and Triplett, E. W. (1998). Trifolitoxin Production Increases Nodulation Competitiveness of Rhizobium etli CE3 under Agricultural Conditions. *Appl Environ Microbiol* 64, 2630–2633.
- Ross-Gillespie, A., Dumas, Z., and Kümmerli, R. (2015). Evolutionary dynamics of interlinked public goods traits: an experimental study of siderophore production in Pseudomonas aeruginosa. *J. Evol. Biol.* 28, 29–39. doi:10.1111/jeb.12559.
- Ruhe, Z. C., Low, D. A., and Hayes, C. S. (2013). Bacterial contact-dependent growth inhibition (CDI). *Trends Microbiol* 21, 230–237. doi:10.1016/j.tim.2013.02.003.
- Rypien, K. L., Ward, J. R., and Azam, F. (2010). Antagonistic interactions among coral-associated bacteria. *Environmental Microbiology* 12, 28–39. doi:10.1111/j.1462-2920.2009.02027.x.

- Sandoz, K. M., Mitzimberg, S. M., and Schuster, M. (2007). Social cheating in Pseudomonas aeruginosa quorum sensing. *PNAS* 104, 15876–15881. doi:10.1073/pnas.0705653104.
- Schink, B. Synergistic interactions in the microbial world. *Antonie Van Leeuwenhoek* 81, 257–261. doi:10.1023/A:1020579004534.
- Schulz, S., and Dickschat, J. S. (2007). Bacterial volatiles: the smell of small organisms. *Nat Prod Rep* 24, 814–842. doi:10.1039/b507392h.
- Shabuer, G., Ishida, K., Pidot, S. J., Roth, M., Dahse, H.-M., and Hertweck, C. (2015). Plant pathogenic anaerobic bacteria use aromatic polyketides to access aerobic territory. *Science* 350, 670–674. doi:10.1126/science.aac9990.
- Shen, B. (2003). Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol* 7, 285–295.
- Smith, S. E., and Smith, F. A. (2011). Roles of arbuscular mycorrhizas in plant nutrition and growth: new paradigms from cellular to ecosystem scales. *Annu Rev Plant Biol* 62, 227–250. doi:10.1146/annurev-arplant-042110-103846.
- Song, C., Schmidt, R., de Jager, V., Krzyzanowska, D., Jongedijk, E., Cankar, K., et al. (2015). Exploring the genomic traits of fungus-feeding bacterial genus Collimonas. *BMC Genomics* 16, 1103. doi:10.1186/s12864-015-2289-3.
- Stintzi, A., Evans, K., Meyer, J. M., and Poole, K. (1998). Quorum-sensing and siderophore biosynthesis in Pseudomonas aeruginosa: lasR/lasI mutants exhibit reduced pyoverdine biosynthesis. *FEMS Microbiol. Lett.* 166, 341–345. doi:10.1111/j.1574-6968.1998.tb13910.x.
- Stoodley, P., K. Sauer, D. G. Davies, and Costerton, J. W. (2002). Biofilms as Complex Differentiated Communities. *Annual Review of Microbiology* 56, 187–209. doi:10.1146/ annu. rev. micro.56.012302.160705.
- Stubbendieck, R. M., Vargas-Bautista, C., and Straight, P. D. (2016). Bacterial Communities: Interactions to Scale. *Front Microbiol* 7, 1234. doi:10.3389/fmicb.2016.01234.
- Suárez-Estrella, F., Arcos-Nievas, M. A., López, M. J., Vargas-García, M. C., and Moreno, J. (2013). Biological control of plant pathogens by microorganisms isolated from agro industrial composts. *Biological Control* 67, 509–515. doi:10.1016/j.biocontrol.2013.10.008.
- Subramanian, S., and Smith, D. L. (2015). Bacteriocins from the rhizosphere microbiome – from an agriculture perspective. *Front Plant Sci* 6. doi:10.3389/fpls.2015.00909.
- Sulieman, S., and Tran, L.-S. P. (2014). Symbiotic Nitrogen Fixation in Legume Nodules: Metabolism and Regulatory Mechanisms. *Int J Mol Sci* 15, 19389–19393. doi:10.3390/ijms151119389.
- Svanbäck, R., and Bolnick, D. I. (2007). Intraspecific competition drives increased resource use diversity within a natural population. *Proc Biol Sci* 274, 839–844. doi:10.1098/rspb.2006.0198.
- Tenaillon, O., Rodríguez-Verdugo, A., Gaut, R. L., McDonald, P., Bennett, A. F., Long, A. D., et al. (2012). The Molecular Diversity of Adaptive Convergence. *Science* 335, 457–461.

doi:10.1126/science.1212986.

- Trapet, P., Avoscan, L., Klinguer, A., Pateyron, S., Citerne, S., Chervin, C., et al. (2016). The Pseudomonas fluorescens Siderophore Pyoverdine Weakens Arabidopsis thaliana Defense in Favor of Growth in Iron-Deficient Conditions. *Plant Physiol.* 171, 675–693. doi:10.1104/pp.15.01537.
- Turner, T. R., James, E. K., and Poole, P. S. (2013). The plant microbiome. *Genome Biol* 14, 209. doi:10.1186/gb-2013-14-6-209.
- Tyc, O., Song, C., Dickschat, J. S., Vos, M., and Garbeva, P. (2017). The Ecological Role of Volatile and Soluble Secondary Metabolites Produced by Soil Bacteria. *Trends Microbiol.* 25, 280–292. doi:10.1016/j.tim.2016.12.002.
- Tyc, O., Zweers, H., de Boer, W., and Garbeva, P. (2015). Volatiles in Inter-Specific Bacterial Interactions. *Front Microbiol* 6, 1412. doi:10.3389/fmicb.2015.01412.
- Van Acker, H., Van Dijck, P., and Coenye, T. (2014). Molecular mechanisms of antimicrobial tolerance and resistance in bacterial and fungal biofilms. *Trends in Microbiology* 22, 326–333. doi:10.1016/j.tim.2014.02.001.
- Velicer, G. J., Kroos, L., and Lenski, R. E. (2000). Developmental cheating in the social bacterium Myxococcus xanthus. *Nature* 404, 598–601. doi:10.1038/35007066.
- Vílchez, J. I., García-Fontana, C., Román-Naranjo, D., González-López, J., and Manzanera, M.(2016). Plant Drought Tolerance Enhancement by Trehalose Production of Desiccation-Tolerant Microorganisms. *Front Microbiol* 7. doi:10.3389/ fmicb.2016.01577.
- Walker, T. S., Bais, H. P., Déziel, E., Schweizer, H. P., Rahme, L. G., Fall, R., et al. (2004). Pseudomonas aeruginosa-plant root interactions. Pathogenicity, biofilm formation, and root exudation. *Plant Physiol.* 134, 320–331. doi:10.1104/pp.103.027888.
- Wandersman, C., and Delepelaire, P. (2004). Bacterial Iron Sources: From Siderophores to Hemophores. *Annual Review of Microbiology* 58, 611–647. doi:10.1146/annurev.micro.58.030603.123811.
- Wei, Z., Yang, T., Friman, V.-P., Xu, Y., Shen, Q., and Jousset, A. (2015). Trophic network architecture of root-associated bacterial communities determines pathogen invasion and plant health. *Nat Commun* 6, 8413. doi:10.1038/ncomms9413.
- Wellman, C. H., Osterloff, P. L., and Mohiuddin, U. (2003). Fragments of the earliest land plants. *Nature* 425, 282–285. doi:10.1038/nature01884.
- West, S. A., Griffin, A. S., and Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evol. Biol.* 20, 415–432. doi:10.1111/j.1420- 9101.2006.01258.x.
- Westerholm, M., Dolfing, J., Sherry, A., Gray, N. D., Head, I. M., and Schnürer, A. (2011). Quantification of syntrophic acetate-oxidizing microbial communities in biogas processes. *Environ Microbiol Rep* 3, 500–505. doi:10.1111/j.1758-2229.2011.00249.x.

- Whipps, J. M. (2001). Microbial interactions and biocontrol in the rhizosphere. *J. Exp. Bot.* 52, 487–511. doi:10.1093/jexbot/52.suppl_1.487.
- Yang, J., Kloepper, J. W., and Ryu, C.-M. (2009). Rhizosphere bacteria help plants tolerate abiotic stress. *Trends Plant Sci.* 14, 1–4. doi:10.1016/j.tplants.2008.10.004.
- Zelezniak, A., Andrejev, S., Ponomarova, O., Mende, D. R., Bork, P., and Patil, K. R. (2015). Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proc. Natl. Acad. Sci. U.S.A.* 112, 6449–6454. doi:10.1073/ pnas.1421834112.

Combining genomic, metabolomic and phenotypic studies to explore the competitive potential of phylogenetically diverse bacteria

II.A. Introduction

Microbes live in dense multispecies communities where ecological interactions are expected to alter their fitness and survival. To secure nutrients and space, bacteria have evolved a panoply of sophisticated mechanisms to compete against opponent cells (reviewed by Ghoul and Mitri, 2016). These mechanisms range from contact-dependent (reviewed in Hayes *et al.,* 2014) to at-distance killing strategies (recently reviewed in Stubbendieck *et al.,* 2016). Antagonistic interactions mediated by diffusible antimicrobials is a well-described competitive mechanism (reviewed in Hibbing *et al.,* 2010, Cornforth and Foster, 2013, Ghoul and Mitri, 2016, Stubbendieck and Straight, 2016, Stubbendieck *et al.,* 2016). Indeed, bacteria utilize a wide arsenal of antimicrobials that have narrow- to broad-spectrum antagonistic activity in order to neutralize closely or distantly related bacterial opponents, respectively (Subramanian and Smith, 2015, Tyc *et al.*,2014,). For instance, bacteriocines are ribosomally synthesized peptides that often target closely related species and have been also reported to inhibit distantly related species (Ghoul *et al.,* 2015, Riley and Gordon, 1999). Alternatively to bacteriocines, bacteria have evolved nonribosomally synthesized antimicrobials that have a board-spectrum antagonistic activity such as nonribosomal peptides (nrps) or polyketides (pks) (Tyc *et al.,* 2017, Traxler *et al.,* 2013). Several soil or plant-associated bacteria secrete specialized metabolites known from the anthropogenic perspective as antibiotics (reviewed by Raaijmakersand Mazzola, 2012, Tyc *et al.,* 2017). Antibiotics have been used since several decades to combat pathogen-mediated diseases. Although the ecological role of antibiotics has been questioned (Davies *et al.,* 2006), these metabolites are still considered as chemical weapons used by bacteria to compete against other microbes (Comforth and Foster, 2015, Abrudan *et al.,* 2015).

Biosynthetic gene clusters (BGCs), which refer to a physically clustered group of genes, encode enzymatic pathways that are necessary for the production of specialized metabolites (originally described as secondary metabolites, Stubbendieck and Straight, 2016). These metabolites have diverse biological functions, including antibiosis such as bacteriocin, nrps or pks (Stubbendieck and Straight, 2016, Madema *et al.*,2015, Cimermancic *et al.*,2014). Democratization of high-throughput sequencing technologies has recently led to an exponential increase in sequenced bacterial genomes. Therefore, *in silico* genome mining for BGCs has become a key upstream methodology to uncover new metabolites and appreciate genomes' diversity in BGCs. The recent sequencing of more than 200 soil and root-associated bacterial genomes offers a unique opportunity to scrutinize BGCs variation and diversity across phylogenetically diverse bacteria that are representative of a host-associated microbiota (*i.e. Arabidopsis thaliana* root-associated microbiota) (Bai *et al.,* 2015). Several computational tools are now available to mine genomes for BGCs such as NaPDoS (Ziemert *et al.,* 2012) and antiSMASH (Madema

et al., 2011, Weber *et al.,* 2015). Few studies have already demonstrated that bacterial genomes harbor a staggering diversity of BGCs (Cimermancic *et al.,* 2014, Donia *et al.,* 2014, Maansson *et al.,* 2016). However, none of these studies focused on analyzing the genomic potential in BGCs of the *Arabidopsis thaliana* root-associated microbiota. Although bacteria harbor a substantial number and diversity of BGCs, only a small proportion of corresponding specialized metabolites are detected downstream through metabolomic analysis. Indeed, most BGCs are known to be silent and not expressed under laboratory conditions (Crüssman *et al.,* 2016, Rutledge and Challis, 2015). Therefore, it is primordial to combine *in silico* genome mining with systematic exploration of bacterial metabolites in order to identify genome-predicted produced antimicrobials. To study bacterial metabolites, nuclear magnetic resonance (NMR) or mass spectrometry (MS) can be performed on different bacterial cultures or supernatants (Krug and Müller, 2014). Both methods detect and structurally elucidate metabolites (Krug and Müller, 2014). The latter method can be coupled with gas chromatography (GC) to detect volatile compounds, or with liquid chromatography (LC) to detect soluble metabolites (Krug and Müller, 2014). Combining genome mining with untargeted metabolomic analysis will undoubtedly provide a more accurate picture of the chemical diversity of metabolites produced by bacteria under conventional laboratory growth and evaluate the number of these metabolites that are genome-predicted. Furthermore, applying both former and latter approaches in the study of soil and root-associated bacteria may reveal the importance of certain specialized metabolites in host-microbe and/or microbe-microbe interactions during the establishment of plant-associated microbiota.

Bacteria that produce antimicrobials have a competitive advantage against antimicrobial-sensitive strains (Gerardin *et al.,* 2016). It is conceivable, therefore, that antimicrobial warfare could be involved in the delineation between ecologically defined bacterial groups (Cordero *et al.,* 2012, Maida *et al.,* 2015). Screen for mutual antagonism can reveal phylogenetic or ecological patterns of inter-bacterial inhibitions and could also serve as basis to assess individually the competitiveness of community members that are ecologically delineated or not. Screen for mutual inhibitions mediated by diffusible antimicrobials have been used often to test the inhibitory activity of host- or environment-associated bacteria against other community members (Cordero *et al.*,2012, Maida *et al.*,2015) or against few indicative strains (Tyc *et al.*,2014, Abrudan *et al.*,2015). In this chapter, I explore the genomic and the metabolomic diversity of soil and *A. thaliana* root-derived bacteria by combining systematic analysis of the bacterial genomes and metabolites. Furthermore, I test the hypothesis that antagonistic interactions between and among culturable and ecologically delineated bacterial groups (*i.e* root-derived bacteria and abundant soil bacteria) are common. The analysis of bacterial genomes revealed that soil and root-derived bacteria harbor different BGCs that encode enzymatic pathway for the biosynthesis of specialized metabolites, including antimicrobials. Some BGC classes are distributed across phylogenetically diverse bacterial strains, whereas other classes are constrained to a defined phylogenetic groups. More importantly, the analysis of genomes revealed that *Streptomycetaceae*, *Nocardiaceae* and *Mycobacteriaceae* isolates are by far the most enriched and diversified in BGCs than

all remaining *Actinobacteria* isolates. The analysis of bacterial metabolites however confirmed that only few of genome-predicted antimicrobials are produced *in vivo* by several bacterial isolates grown on synthetic medium. The goal of this study is to define two groups of bacteria with contrasting competitive potential. The screen for mutual inhibitions did not only provide a ground for choosing highly competitive and highly sensitive community members, but also revealed interesting phylogenetic and ecological patterns in inter-bacterial competitive interactions. Indeed, screen for inter-bacterial antagonistic interactions revealed that most of *Actinobacteria* isolates are sensitive to produced antimicrobials and revealed a fierce competition between abundant soil bacteria and root-derived bacteria. Combination of multi-faceted systematic analyses of the bacterial competitiveness that explore the genomes, metabolomes and "antagonome" are important to define highly competitive and highly sensitive bacteria and ultimately understand the fundamental role of antagonistic interactions in altering the community structure of host-associated microbiota.

II.B. Mining the bacterial genomes reveals phylogenetic patterns in the distribution of biosynthetic gene clusters

Bacterial genomes harbor a wide diversity of biosynthetic gene clusters (BGCs) that encode for enzymatic pathways involved in the production of specialized metabolites. These metabolites consist of diverse classes of molecules such as polyketides (pks), nonribosomal peptides (nrps), bacteriocins, terpene and many other classes. The identification of BGCs has become possible thanks to the development of powerful predictive *in silico* tools such as the antibiotics and Secondary Metabolite Analysis Shell scripts (antiSMASH) (Madema *et al.*, 2011). In order to explore the genomes for BGCs, I submitted 198 bacterial draft genomes to antiSMASH. The corresponding bacterial strains have been isolated from *Arabidopsis thaliana* roots or from Cologne agricultural soil and represent two ecologically delineated groups; root-derived bacteria or abundant soil bacteria, respectively (further details indicated in Materials and Methods). Abundant soil bacteria represent 15% of the analyzed genomes and cover three main phyla (*Actinobacteria*, *Proteobacteria* and *Firmicutes*). The analysis of 198 genomes uncovered 30 BGCs classes (out of 40 known and detectable by antiSMASH) for a total of 1,404 clusters. In descending order, nrps, terpene and bacteriocin are the most abundant classes (**Figure 1**). Interestingly, the BGC class other, which refers to unknown biosynthetic gene cluster, represents the top fourth most abundant class. These data indicate not only that the bacterial strains harbor known BGCs, but also harbor cryptic BGCs with potentially novel specialized metabolites.

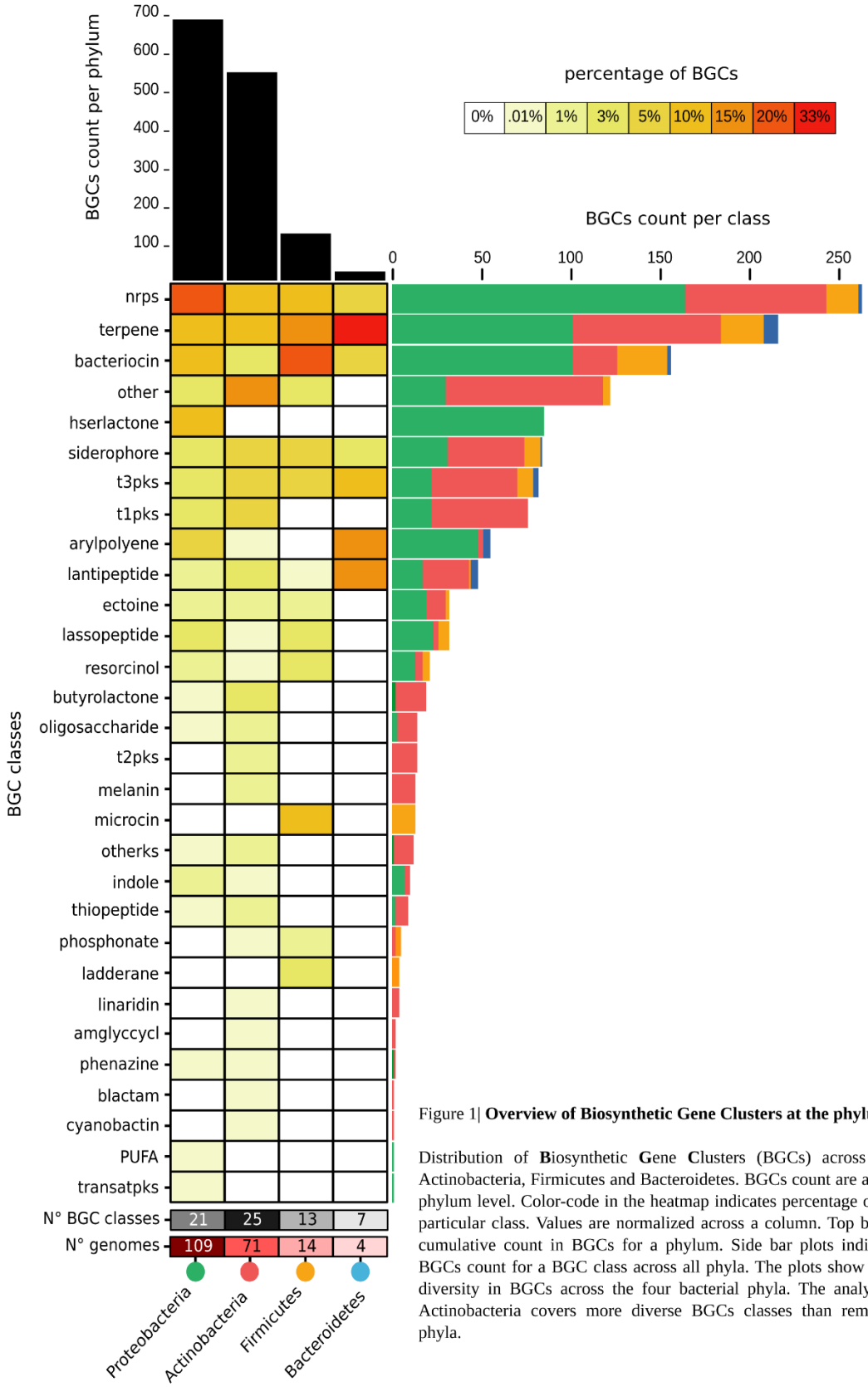


Figure 1| **Overview of Biosynthetic Gene Clusters at the phylum level.**

Distribution of **B**iosynthetic **G**ene **C**lusters (BGCs) across Proteobacteria, Actinobacteria, Firmicutes and Bacteroidetes. BGCs count are aggregated at the phylum level. Color-code in the heatmap indicates percentage of BGCs for one particular class. Values are normalized across a column. Top bar plots refer to cumulative count in BGCs for a phylum. Side bar plots indicate cumulative BGCs count for a BGC class across all phyla. The plots show distribution and diversity in BGCs across the four bacterial phyla. The analysis reveals that Actinobacteria covers more diverse BGCs classes than remaining bacterial phyla.

In the set of bacterial genomes used in this analysis, 55% represent *Proteobacteria,* whereas *Actinobacteria*, *Firmicutes* and *Bacteroidetes* represent 36%, 7% and 2% of analyzed genomes, respectively. The analysis of bacterial genomes reveals that only few BGC classes are predicted across all phyla and they are nrps, terpene, bacteriocin, siderophore and t3pks. These classes are known to include antimicrobials or metabolites employed during competitive interactions (Tyc *et al.,* 2017). Further analysis revealed that these common BGC classes are not evenly abundant across all phyla (**Figure 1**). For instance, the class nrps is the most abundant BGC class in *Proteobacteria*, whereas the class other is the most abundant BGC class in *Actinobacteria* (**Figure 1**). These data are indicative that certain BGC classes might be preferentially enriched within a particular taxonomic group. However for further assertion, it is important to include more representative genomes for *Firmicutes* and *Bacteroidetes*, since the number of representative genomes for each phylum is largely disproportionate between phyla. Regarding BGCs diversity, more diverse classes are detected in *Actinobacteria* than in *Proteobacteria,* although the latter phylum is represented by 109 genomes and 694 BGCs are predicted from these genomes (**Figure 1**). In contrast to BGC classes predicted across all bacterial phyla, few classes appear to be uniquely predicted in one phylum. For example, the class microcin is only predicted in *Firmicutes*, hserlactone (for homoserine lactone) detected only in *Proteobacteria* or the class t2pks is only found in *Actinobacteria* (**Figure 1**). The analysis of bacterial genomes at the phylum level indicates that BGCs differ qualitatively and quantitatively and these difference cannot be explained only by a bias in representative genomes accfor each phylum. Moreover, this analysis indicates that certain BGC classes are widely predicted across all phyla and other classes are rather phylogenetically constrained which suggest that bacteria employ similar or different classes of specialized metabolites to mediate host-microbe or microbe-microbe interactions.

Since the analyzed genomes cover 25 bacterial families, it is possible to explore commonness and uniqueness in BGC classes at the family level. Each bacterial phylum in the set of genomes is represented by at least two families, except for *Bacteroidetes* that is represented by one family. Although common BGC features are observed at the phylum level, it does not systematically indicate that all bacterial families of a same phylum harbor that particular BGC class. Therefore, I examined BGCs distribution and diversity across all bacterial families (**Figure 2**). The analysis of BGCs at the family level indicates that *Streptomycetaceae* is by far the most enriched (315 gene clusters) and most diversified (24 different classes) in BGCs (**Figure 2**). Interestingly, this analysis indicates that all remaining bacterial family, except *Streptomycetaceae*, hardly exceed 173 clusters or 12 classes (**Figure 2**). This finding corroborate the longstanding acceptance that *Streptomycetaceae* has the genetic potential for being prolific producer of specialized metabolites (Rutledge and Challis, 2015). Strikingly, *Streptomycetaceae* appears as an exception within *Actinobacteria* since top second, third and fourth most abundant bacterial families in BGCs belong to *Proteobacteria* (*Rhizobiaceae*, *Xanthomonadaceae* and *Comamonadaceae*, respectively). *Mycobacteriaceae* also appears as promising family, but more representative genomes are needed to ascertain this observation (**Figure 2**). All bacterial families, except

the top fourth indicated above, show less than 100 BGCs, however their richness in BGC classes remain variable (**Figure 2**). More precisely, the bacterial families harbor between 2 to 12 BGCs classes and between 3 to 173 BGCs, excluding *Streptomycetaceae* (**Figure 2**). Importantly, several of the bacterial families represented by less than 10 genomes and with less than 100 BGCs are predicted to harbor more than 10 different BGCs classes. For instance, *Xanthomonadaceae* is predicted with 116 clusters that cover only 9 different classes. In the contrary, several other bacterial families cover more than 10 classes with less than 100 total BGCs count (**Figure 2**). Remarkably, the only representative of *Nocardiaceae* is predicted to harbor 10 different BGC classes for only 30 BGCs (**Figure 2**). The analysis of BGCs at the family level suggests that quantitative and qualitative differences are observable at the family level and these differences could not be only explained by a bias in representative genomes between bacterial families, but also reflect phylogeny- and/or ecology-driven diversity between bacteria. More importantly, this analysis highlights that *Streptomycetaceae* harbors a remarkable number and diversity of BGCs indicating that these bacteria have evolved diverse metabolites that mediate microbe-microbe or microbe-host interactions.

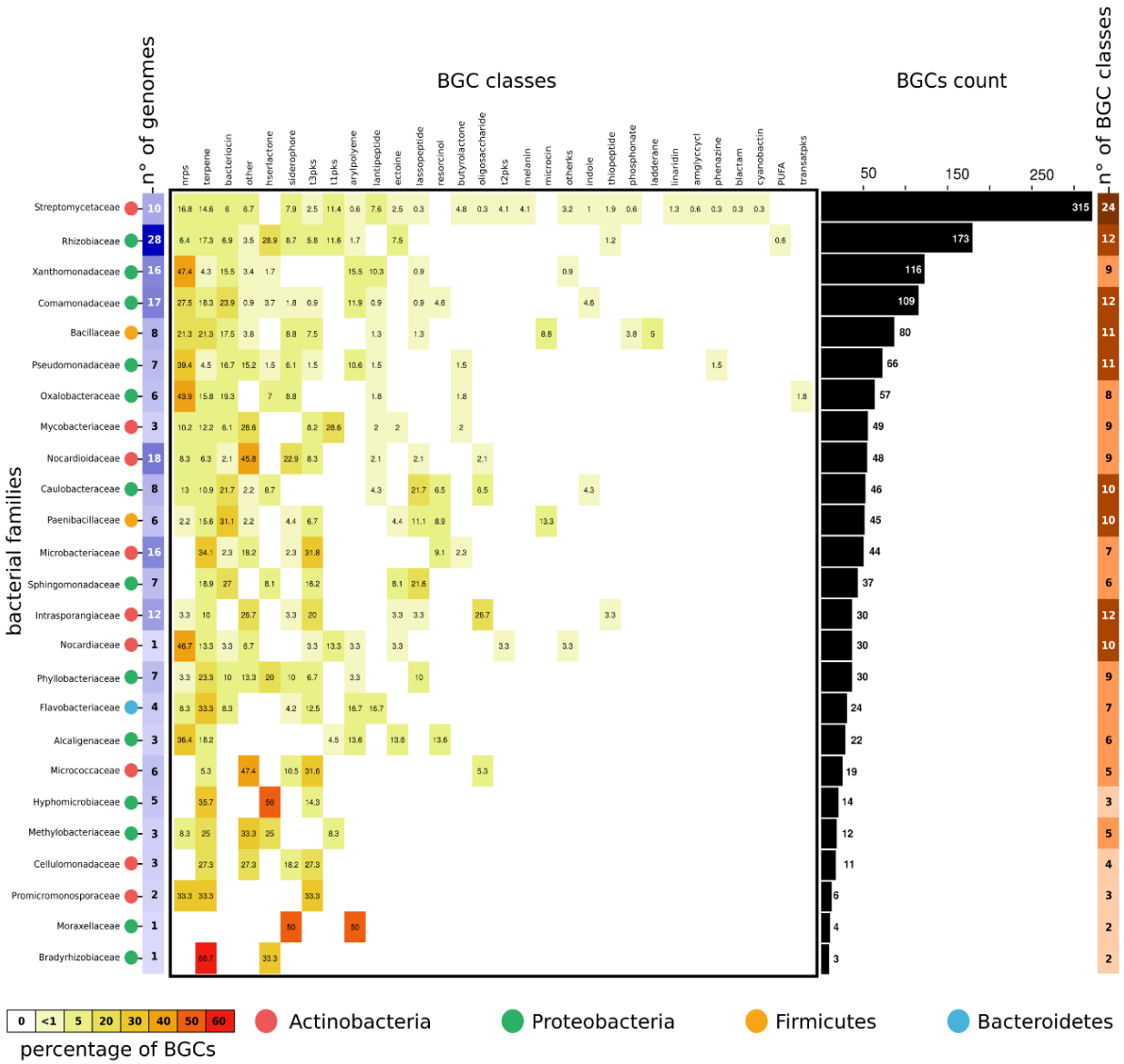


Figure 2| **Distribution in BGCs across the bacterial families.**

The heatmap indicates distribution in BGCs across the 25 bacterial families. Indicated values are normalized by the sum of BGCs for each corresponding family. Bacterial families are indicated in y-axis and BGCs classes are represented in x-axis. Color code in circular shape refers to the phylum. Each family is represent by one to several genomes that is indicated by the column n° of genomes. Bar-plots indicate cumulative count in BGCs per bacterial family. The column n° of BGC classes indicates BGC diversity. This analysis reveals qualitative and quantitative differences in BGCs at family level and highlights the enrichment in BGCs in *Streptomycetaceae*.

Although differences are observed in BGCs between bacterial families, there are few BGCs classes that are predicted across phylogenetically diverse bacterial families. For example, the class terpene is predicted across almost all the bacterial families, except in *Moraxellaceae* that is represented by one genome (**Figure 2**). Terpenes are volatile compounds that fulfill diverse biological activities, including antagonistic interactions (Tyc *et al.*, 2017). Interestingly, almost all bacterial families belonging to *Proteobacteria* (except *Moraxellaceae* and *Alcaligenaceae* represented by three isolates) are predicted to harbor BGCs involved in the biosynthesis of hserlactones, molecules employed by many bacteria for quorum sensing (Federle and Bassler., 2003). All families belonging to *Actinobacteria*, *Firmicutes* and *Bacteroidetes* are predicted to possess t3pks, whereas several families within *Proteobacteria* lack this class (**Figure 2**). The Few bacterial families within *Proteobacteria* that are predicted to harbor t3pks

belong two α-*Proteobacteria,* two β-*Proteobacteria* and one γ-*Proteobacteria* families. In contrast to BGCs that are represented across phylogenetic diverse strains, the class microcin is only detected in the two bacterial families representing *Firmicutes* (*Bacillaceae* and *Paenibacillaceae*) (**Figure 2**). Microcins are small ribosomally synthesized low-molecular-mass antibacterial molecules that have narrow-spectrum antagonistic activity (Duquesne *et al.*,2007) and have been reported to mediate *in vivo* intra-specific (between *Entrobacteriaceae* bacteria) competitive interactions (Sassone-Corsi *et al.*, 2016). The analysis of bacterial genomes at the family level indicates that very few BGC classes (like terpene, nrps, pks) are predicted across all surveyed genomes. More importantly, this analysis highlights the heterogeneity in BGCs across bacterial families belonging to a same phylum.

To further reveal BGCs diversity in the genomes, I performed an additional analysis at the isolate level and the data are depicted in **Figure 3** and **Supplementary Figure 2**. The set of the genomes covers 25 bacterial families. Most bacterial families are represented by multiple genomes, except four families that are represented by one genome. From 198 bacterial genomes, ~ 83% (164 genomes) of the isolates possess less than 10 BGCs and ~ 94% (186 genomes) of them are predicted to harbor less than 10 different BGCs classes (**Figure 3**, **supplementary Figure 2**). Impressively, 80% of *Actinobacteria* isolates show very low BGC diversity (<5 BGCs), whereas the remaining 20% encode a staggering diversity of BGCs (>14 BGCs) (**Figure 3**). The bacterial isolates that contribute most to this striking BGC expansion belong to three bacterial families: *Streptomycetaceae, Nocardiaceae* and *Mycobacteriaceae* (**Figure 3**). Notably, *Streptomycetaceae* isolates are predicted to harbor on average ~31±8 BGCs and ~14±2 BGCs classes. The only representative of *Nocardiaceae* harbors 30 BGCs spanning across 10 different BGCs classes, whereas the three *Mycobacteriaceae* isolates contain in average 16±3 clusters belonging to 7±1 classes. This observation is unique in a way that there is no similar expansion observed within *Proteobacteria, Bacteroidetes* or *Firmicutes* isolates used in this study. It is also important to highlight that the expansion contained in BGCs within *Streptomycetaceae, Nocardiaceae* and *Mycobacteriaceae* could not be explained by the genome size since there is no correlation between genome size and number of predicted BGCs **(Supplementary Figure 1)**. Moreover, only 20% of *Bacteroidetes*, *Firmicutes* and *Proteobacteria* isolates show less than 5 BGCs against 99% of *Actinobacteria* isolates, if *Streptomycetaceae*, *Mycobacteriaceae* and *Nocardiaceae* are not counted (**Figure 3**). Interestingly, this observation is in discrepancy with Cimermancic *et al.*,2014 that have reported a linear trend between BGCs count and genome size. The analysis of BGCs at the isolate level highlights two main features; only very limited and taxonomically constrained *Actinobacteria* isolates are highly enriched and diversified in BGCs, most of *Actinobacteria* isolates, except in *Streptomycetaceae, Mycobacteriaceae* and *Nocardiaceae,* are very limited in BGCs number and diversity. In a more specific manner, these data also reveal that several *Streptomyces* isolates possess BGCs classes that are not commonly predicted in other bacterial isolates. For instance, butyrolactone quorum sensing-like molecule for activation of antibiotic production (Takano, 2006) or "t2pks" known to synthesize a wider range of bioactive molecules that are clinically useful (Hertweck *et al.,* 2006) are

almost exclusive to *Streptomycetaceae* (**Supplementary Figure 2**). Overall, several biosynthetic classes are rather very rare (less than 5% of coverage) to uncommon (between 15%-25% of coverage) than very common (supp. to 75% of coverage) among bacterial isolates (**Supplementary Figure 2**). On the other hand, most abundant BGCs class is terpene, which is predicted in 147/198 of the bacterial genomes (**Supplementary Figure 2**). As highlighted earlier, terpenes are a class of volatile metabolites that have diverse biological functions including antibiosis (Tyc *et al.*, 2017). Since terpenes are volatile molecules, this indicates that microbe-host or microbe-microbe interactions can be mediated from very long distances. The class bacteriocin is also often predicted (94/198) from the genomes, more frequent in *Firmicutes*, *Proteobacteria* and *Bacteroidetes* (**Supplementary Figure 2**). Within *Actinobacteria*, only *Streptomycetaceae*, *Nocardiaceae* and *Mycobacteriaceae* isolates are predicted to produce bacteriocines. These data indicate that several *Actinobacteria* isolates that do not belong to *Streptomycetaceae*, *Nocardiaceae* and *Mycobacteriaceae* are poorly enriched in BGCs. It is therefore plausible to hypothesize that during competitive interactions, several *Actinobacteria* isolates could be be strongly inhibited by other community members. Although certain BGCs follow restrained phylogeny, this analysis failed to provide any meaningful differences between the two ecologically-delineated bacterial groups, root-derived and abundant soil bacteria (**Supplementary Figure 2 and 3**). Indeed, neither qualitative nor quantitative differences in BGCs are observable across the bacterial isolates that recapitulate the two bacterial populations from these two distinct compartments. In conclusion, soil and root-associated bacterial microbiota members harbor different classes of BGC that encode for enzymatic pathways responsible for the synthesis of specialized metabolites including antimicrobial metabolites. Interestingly, qualitative and quantitative differences in BGCs are commonly observable across tested strains. Lateral gene transfer might be a mechanism that contributes to the observed heterogeneity in BGCs, as other mechanisms. More importantly, the analysis of genomes clearly shows that *Streptomycetaceae*, *Nocardiaceae* and *Mycobacteriaceae* isolates are exceptionally diversified and enriched in BGCs compared to other *Actinobacteria* isolates. These observed differences in BGCs might impact a species' competitiveness and are indicative of a strain genetic potential for the biosynthesis of specialized metabolites.

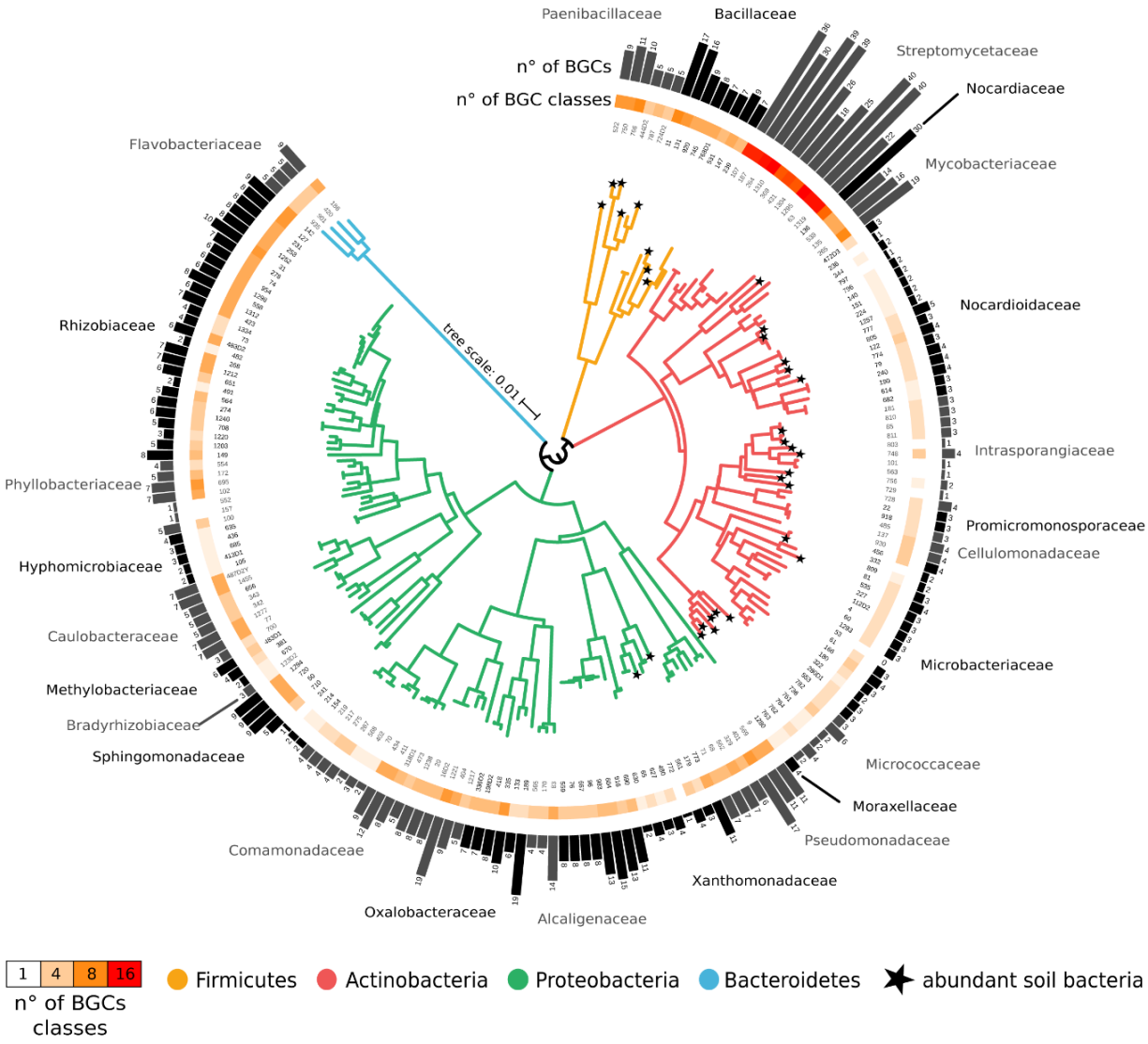


Figure 3| **Distribution of BGCs across phylogenetically diverse bacterial strains.**

The circular phylogenetic tree is based on full length 16s rRNA gene and depicts the number of BGCs and BGC classes across the 198 bacterial strains. Color code in branches indicates the phylum. Abundant soil bacteria are indicated by a star. Black and gray colors in the bar-plots indicate bacterial families. The analysis of BGCs at the strain level reveals that *Streptomyces*, *Nocardiaceae* and *Mycobacteriaceae* strains are enriched in BGCs.

II.C Plant-associated and soil-derived bacteria produce antimicrobial molecules

The genome mining for BGCs of phylogenetically diverse bacteria revealed that root-derived and abundant soil bacteria both harbor biosynthetic capabilities to produce specialized metabolites including known antimicrobial molecules. Although bacteria are predicted to harbor BGCs encoding for antimicrobials, these clusters can be silent and not expressed under laboratory conditions (Rutledge and Challis, 2015). It is therefore crucial to determine whether some of the predicted antimicrobials are produced *in vivo* in synthetic growth medium. To systematically analyze bacterial metabolites and identify known secreted antimicrobial molecules, we profiled the metabolites produced by each of the 198 bacterial strains using liquid chromatography tandem spectrometry (HPLC-MS/MS). The metabolites are extracted from separately grown bacteria on 25% tryptic soy agar, using two organic solvents with different polarities, ethyl acetate and methanol. The analysis of 396 samples by HPLC-MS/MS generated around 200,000 mass spectra that are analyzed through the Global Natural Products Social (GNPS) molecular networking work flow (Guthals *et al.,* 2012). The spectra of minimum of four fragment ions and with at least two identical spectra are merged into a consensus node. The molecular networking algorithm generated a network with 3,316 nodes, representing different parent masses, after removing nodes affiliated to non-inoculated agar. In order to compare metabolite production in the four bacterial phyla, nodes are aggregated at the phylum level. This analysis shows that 2,669 and 2,250 nodes derived from *Proteobacteria* and *Actinobacteria* isolates, respectively, whereas *Firmicutes* and *Bacteroidetes* show less than 1,000 nodes (**Figure 4, panel-A**). Although there is a taxonomic bias in representative isolates across the four phyla, our analysis demonstrates that a substantial proportion of nodes are overlapping between two phyla (**Figure 4, panel-B**). To explore uniqueness and commonness of the identified nodes, we calculated the number of nodes that are unique to each phylum or shared between different phyla. The analysis reveals that ~44% of recovered nodes are uniquely shared between *Proteobacteria* and *Actinobacteria,* but for instance only one node is found to be uniquely shared between *Firmicutes* and *Bacteroidetes* (**Figure 4, panel-B**). Notably, *Proteobacteria* and *Actinobacteria* show a large fraction of unique nodes (830 and 444, respectively, **Figure 4, panel-B**). *Firmicutes* and *Bacteroidetes* show less than 100 unique nodes per phylum, which reflect restricted number of profiled isolates (14 and 4, respectively, **Figure 4, panel-B**). These data nonetheless suggest that each phylum has a unique chemical signature. However, it is still conceivable that a member of a phylum do not produce all unique nodes that are specific to that phylum. Moreover, since a large proportion (>75%) of nodes are shared between at least two members belonging to two different phyla, it is expected that bacteria belonging to a same phylum would not systemically cluster together according to their chemical profiles.

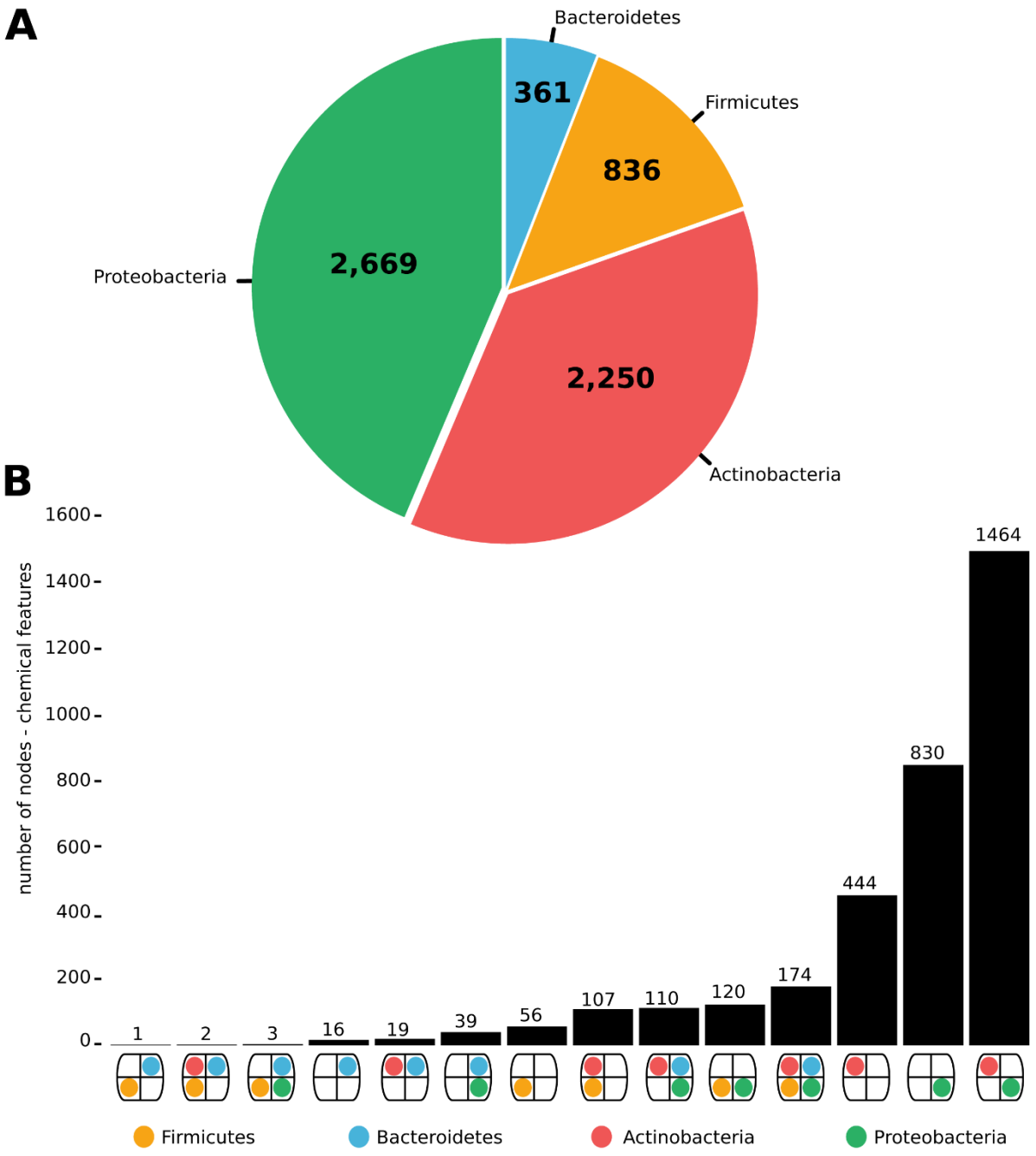


Figure 4| **The analysis of produced bacterial metabolites by LC-MS/MS reveals percentages of shared and unique chemical features across bacterial phyla.**

The graphs show the LC-MS/MS analysis at the phylum level of organic extracts obtained from 198 isolates grown separately in 25% tryptic soy agar. **Panel-A**, the pie chart depicts absolute number of retrieved chemical features (nodes) from secreted metabolites of bacterial strains that are aggregated at phylum level. Large proportion of the chemical features are attributed to Proteobacteria and Actinobacteria. **Panel-B**, the bars-plots show chemical features that are unique to one phylum or shared by at least two bacterial phyla. 174 chemical features are shared by all phyla.

To explore, further, the chemical relatedness between all isolates, I performed PCo analysis using Sørensen index on binary (presence/absence) node profiles of all isolates. The ordination analysis shows that bacteria belonging to the same phylum or family do not systematically cluster together in the chemical space (**Figure 5**). This analysis suggests that phylogenetically-related strains can produce remarkably different blend of metabolites (**Figure 5**). Counter intuitively, several bacterial isolates

belonging to different phyla could cluster close to each other in the chemical space. This observation indicates that although an early evolutionary divergence between bacterial isolates belonging to different phyla, the bacteria can still produce metabolites with high resemblance chemistry (**Figure 5**). Our data support the hypothesis that phylogenetic signal at the phylum level is not enough to discriminate between bacterial isolates in the chemical space and indicates a heterogeneity in the chemistry of secreted bacterial metabolites. These data join the heterogeneity in BGCs of analyzed bacterial genomes and point out that bacteria diversify their competitive arsenal dynamically through horizontal gene transfer or diversification in biosynthetic modules through other evolutive processes (Madema *et al.*, 2014).

Although the general trend indicates clearly a substantial heterogeneity in secreted metabolites, few exceptions illustrate the opposite trend. From the PCo analysis, it is clear that few strains tend to cluster close to their family congeners, as illustrated for instance by members from *Nocardioidaceae* or *Comamonadaceae* (**Figure 5**). Interestingly, the ordination analysis also reveals that except two isolates, all soil abundant bacteria cluster on the left side of the graph (**Figure 5**). However, more sampling depth for abundant soil bacteria is needed prior to ascertain this observation. To further investigate features that can discriminate between bacterial isolates, we recovered nodes that are unique to one isolate. From 198 isolates, approximately 33% do not show any unique chemical feature (**Figure 6**). Interestingly, 24% of the isolates possess at least a single unique node and more than 80 bacterial strains show more than one unique node (**Figure 6**). On another aspect, this analysis reveals that only few isolates have more than 10 unique nodes and unexpectedly the isolate N°335 belonging to the genus *Massilia* have 68 unique nodes (**Figure 6**). These data indicate that a large proportion of isolates have unique chemical signatures and only few of them are enriched in unique nodes. All considered, unique nodes revealed in this analysis are chemical fingerprint for the bacteria that reflect their genomic heterogeneity.

Additionally to the metabolomic-comparative analysis, the study of bacterial metabolites is also aimed to identify bioactive molecules. GNPS analysis identified 247 molecular families from 3,316 nodes. Only four known antimicrobial molecules could be detected from the retrieved metabolites (**Figure 7**). For instance, phenazine and brabantamide are two antibiotics exclusively produced by two *Pseudomonas* isolates n°569 or n°401, respectively (**Figure 7**). Biosynthetic gene cluster encoding for the enzymatic pathway responsible for the production of phenazine have been *in silico* predicted in these isolates. Phenazines are metabolites with potent antimicrobial activity (Borrero *et al.*, 2014). Similarly, nactines are antibiotics specifically detected from the metabolites of two *Streptomyces* isolates (n°63 and n°1295, **Figure 7**). Nactines are polyketide ionophore antibiotics known to alter the sterochimistry of K^+, Na^+ or NH_4^+ of opponent cells and mainly are secreted by *Streptomyces* strains (Wang *et al.*, 2014, Kusche *et al.*, 2009). In contrast to above-indicated antibiotics that are phylogenetically constrained, several bacterial isolates that are phylogenetically diverse (n°11, 65, 73, 85, 135, 166, 172, 227, 553, 559, 568, 627, 729, 736, 766, 782, 811, 954, 1277, 198D2, 472D3, 483D1, and 768D1) produce cyclic dipeptides (**Figure 7**). Cyclic dipeptides are known antibiotic secreted by diverse bacteria

(Li *et al.*, 2008) but also have known to have broad biological activity (Abbamondi *et al.*, 2014). It is important to highlight here that the biosynthetic gene clusters for corresponding above-indicated antibiotics have been *in silico* predicted from the bacterial genomes. It is therefore unlikely that the retrieved antibiotics are derivative from the medium.

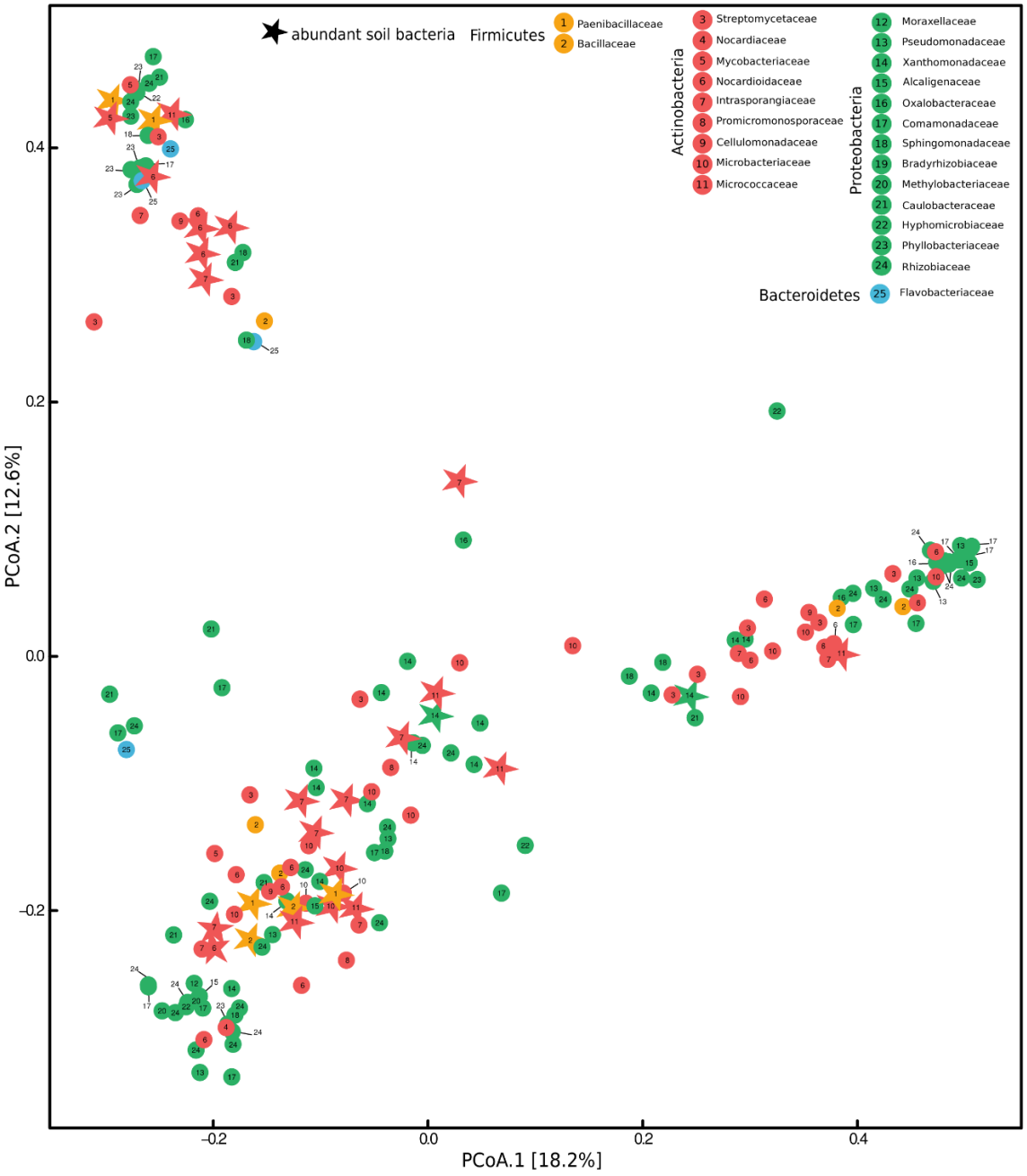


Figure 5| **Bacteria show substantial heterogeneity in the secreted metabolites.**

The graph shows principal component analysis of bacterial metabolomic profiles. Each shape in the graph corresponds to one bacterial isolate. The color code indicates bacterial phylum and star shape indicates abundant soil bacteria. Number within a shape refers to a bacterial family. Distances between isolates in the chemical space correspond to Sørensen index computed on binary (presence/absence) node profiles between all isolates. PCo analysis shows that phylum is not sufficient to distinguish between the chemistry of the isolates and that bacteria are rather spread in the chemical space.

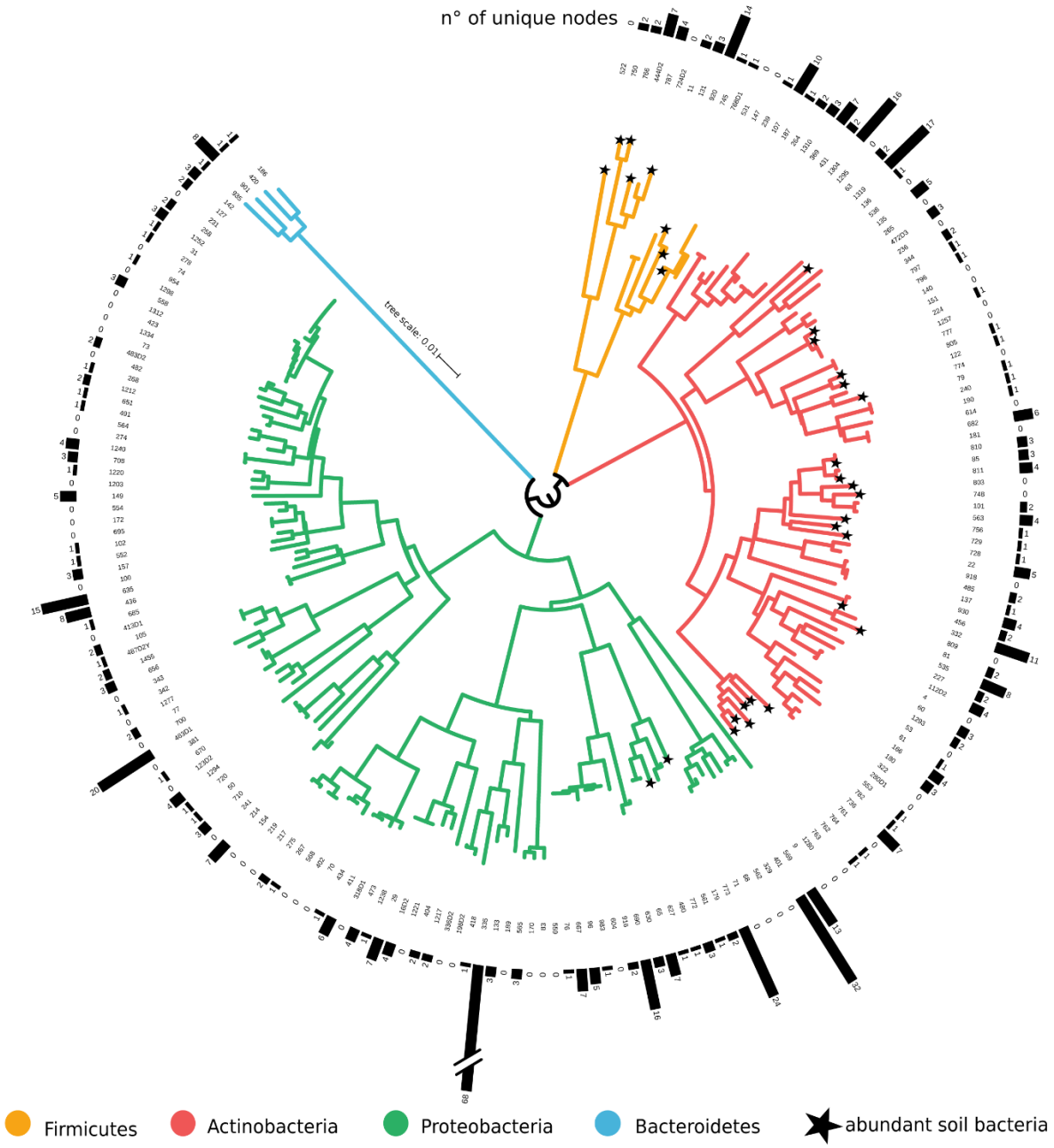


Figure 6| **Several bacterial isolates possess unique chemical signature**

The circular phylogenetic tree depicts 16s rRNA gene based phylogeny of the bacterial strains. Colors in tree branches refer to phyla. Stars indicate abundant soil bacteria. Bar-plots show number of unique nodes retrieved form each single isolate. Most of isolates show a unique chemical signature.

The limited number of detected antimicrobials suggests that 1- most of antimicrobials are not produced in sufficient quantity to be detected by our method, 2- these metabolites are not produced under the applied conditions or 3- the *in silico* predicted BGCs encoding these antimicrobials are not functional. Since most of the identified nodes are unknown, our results suggest that soil and plant-associated bacteria possess the potential for novel specialized metabolites with unknown biological activities. It is also not exclude that bacteria secrete antibiotic upon interactions with competitor cells. Collectively, these data indicate that the bacteria possess chemical weapons that mediate inter-strains antagonistic interactions in order to compete against other community members.

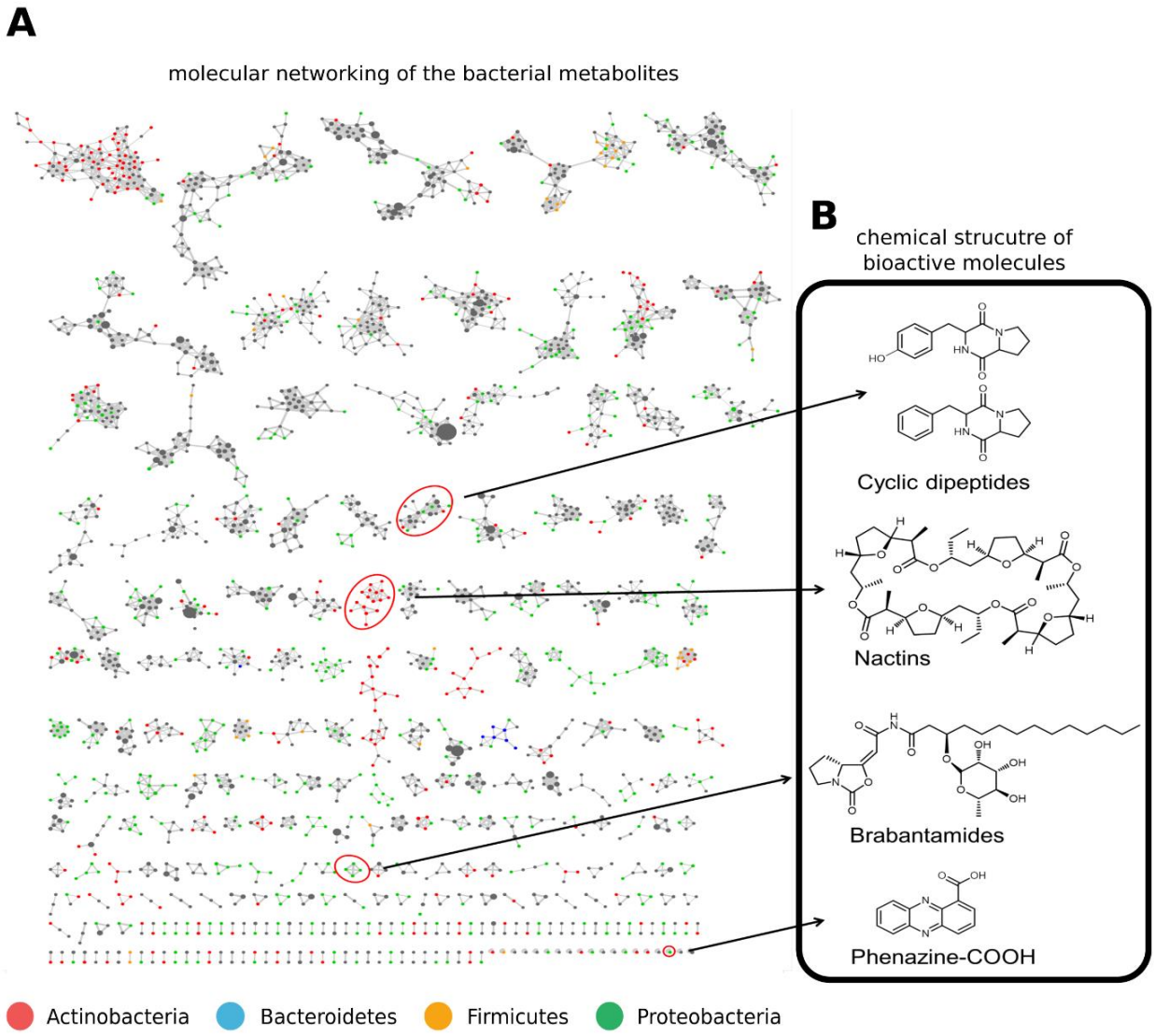


Figure 7| ***Arabidopsis* root-derived and abundant soil bacteria produce antimicrobial molecules**

Panel-A, molecular network of organic extracts from 198 bacterial isolates grown in 25% tryptic soy agar. The network consists of 3,314 nodes and 247 clusters with at least two nodes. Shared metabolites between at least two phyla are indicated by gray color. Nodes that are unique to one phylum are colored according to the phylum color-code. Node size corresponds to number of obtained spectra and indicates the abundance of the retrieved chemical feature. **Panel-B**, chemical structure of bacterially produced antimicrobial molecules. Phenazines and Brabantamides or Nactins are exclusively produced by Proteobacteria or Actinobacteria isolates, respectively. Cyclic dipeptides are produced by several isolates belonging to different phyla.

II.D. Most *Actinobacteria* strains are highly sensitive to antimicrobials secreted by community members

The antagonistic bacteria-bacteria interactions assay. The systematic analysis of the bacterial genomes and metabolomes revealed that soil and root-associated bacteria harbor diverse BGCs and produce a wide range of metabolites, likely conferring a competitive advantage through antimicrobials-mediated antagonistic interactions. Diffusion of antagonistic molecules may play an important ecological role for bacteria to prevent territoriality's invasion by closely or distantly related species or to compete against ecologically delineated bacteria. It is therefore conceivable that bacteria engage in intra-specific (*i.e.* within root-derived bacterial group and within abundant soil bacterial group) and/or inter-specific (*i.e.* between root-derived and abundant soil bacteria) antagonistic interactions. In this section, I test the hypothesis that antagonistic interactions mediated by the production of antimicrobials are observable within and between root-derived and abundant soil bacteria. Moreover, since both *in silico* and metabolomic analyses showed that each strain has different set of genetic and chemical signatures, I also test the hypothesis that competitive antagonism is not equally effective across the isolates, but rather few members have higher competitive potential. In order to test both former and latter hypotheses, I developed an antagonistic bacteria-bacteria interactions assay (ABBA). The screen consists of spotting several bacterial isolates on top of a lawn bacterium. Spot bacteria are referred as producers since they can produce diffusible antimicrobials that inhibit the growth of the lawn bacterium. The lawn bacterium is referred by target strain. A bacterial lawn corresponds to a confluent cellular growth on top of the agar and visible to the naked-eye. An antagonistic interaction is materialized by the presence of a zone of clearance that is known as a halo of inhibition. The halo of inhibition is indicative of the secretion of diffusible bactericidal molecules by a spot bacterium that inhibit the growth of the lawn bacterium. In this precise case, the lawn bacterium is sensitive to the antagonistic activity of the spot bacterium and the spot bacterium exert an antagonistic activity against the lawn bacterium (further information are provided in Materials and Methods). The halo of inhibition can vary in size and is dependent on both the producer and the target strains.

Bacteria engage in antagonistic interactions. Using the ABBA screen, I tested mutual inhibitions mediated by diffusible compounds for all 198 isolates. Over 39,204 tested interactions (198 producers against 198 targets), 1,011 (2.5%) of the interactions show a halo of inhibition visible to the naked-eye (an example is illustrated in Materials and Methods) (**Figure 8**, **panel-A**), suggesting that growth inhibition mediated by the secretion of antimicrobials occurs relatively rarely under the tested conditions. However, it is important to note that 66% of the bacterial isolates could inhibit at least one bacterium and that 62% of the isolates were at least inhibited once (**Figure 8**, **panel-B**). Therefore, both former and latter observations suggest that a large proportion of the bacterial isolates are able to engage in antagonistic interactions through production and secretion of diffusible antimicrobials. Interestingly, not a single isolate is sensitive to all other bacteria or show extensive inhibitory frequencies. Indeed, both sensitivity and antagonistic activity do not exceed 20% and 25% of all tested interactions for any

single bacterium, respectively (**Figure 9**). Moreover, bacterial isolates that are inhibited more than 10 times concern only 20% of the population and isolates that inhibit more than 10 other isolates do not exceed 19% of the population (**Figure 9** and **Supplementary Figure 4**). Thus, most of the isolates antagonize or are inhibited between 1 to 10 times (**Figure 9** and **Supplementary Figure 4**). Taken together, these data not only indicate that bacteria are able to inhibit other bacterial isolates through the secretion of antimicrobials but also resistance to antagonistic activity is a common feature among root-derived and abundant soil bacteria.

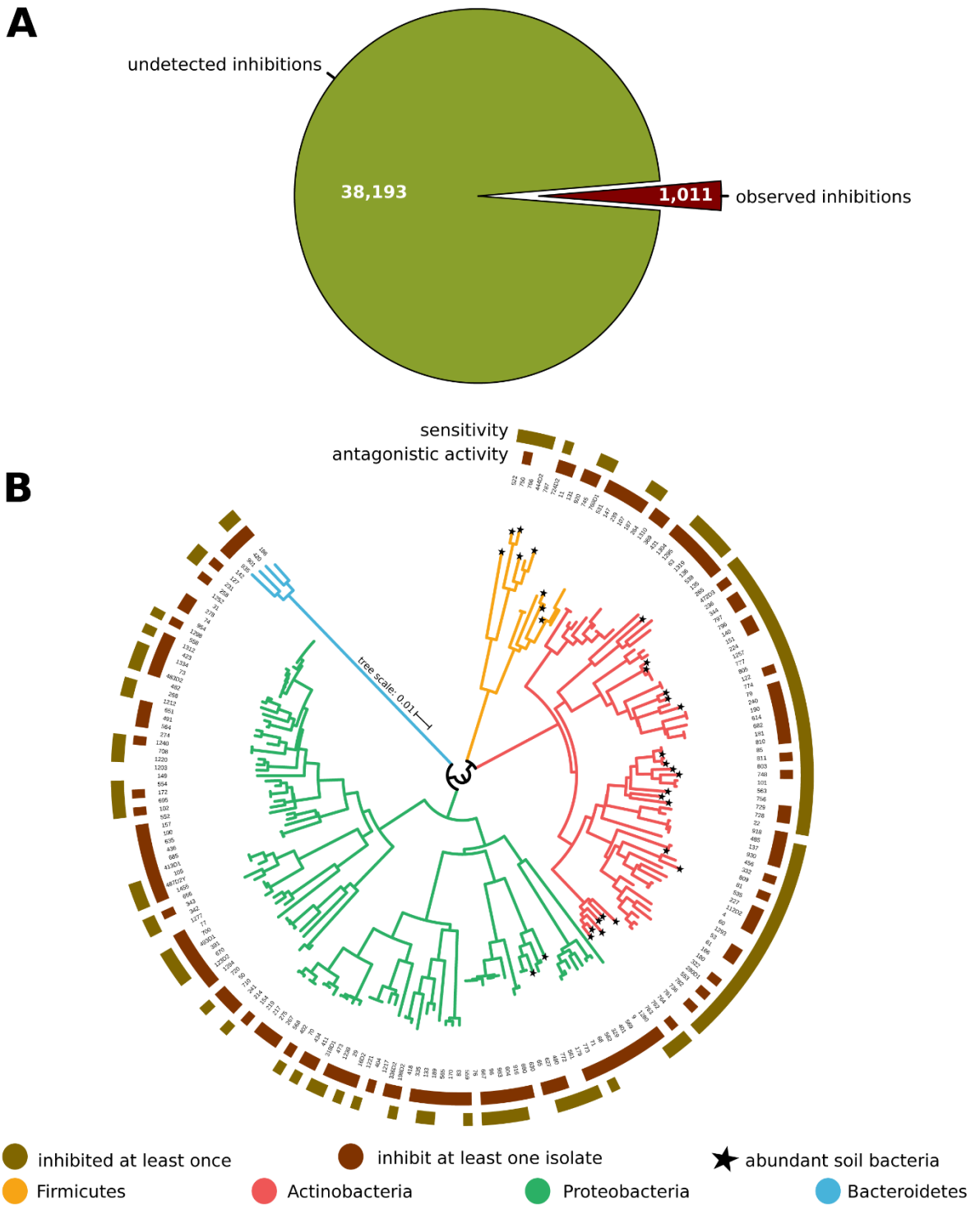


Figure 8| **Several bacterial strains engage in antimicrobials-mediated antagonistic interactions**

Panel-A, the pie chart depicts percentage of observed inhibitions after 39,204 interactions haveb tested. 2.5% of overall interactions show a halo of inhibition that is indicative of antagonistic interaction between two bacterial strains. **Panel-B**, the phylogenetic tree portrays bacterial isolates that are inhibited by at least one other community member or that inhibit at least one other community member. 66% of the bacterial isolates inhibit at least one other bacterial strains and 62% of the isolates are sensitive to at least one bacterium.

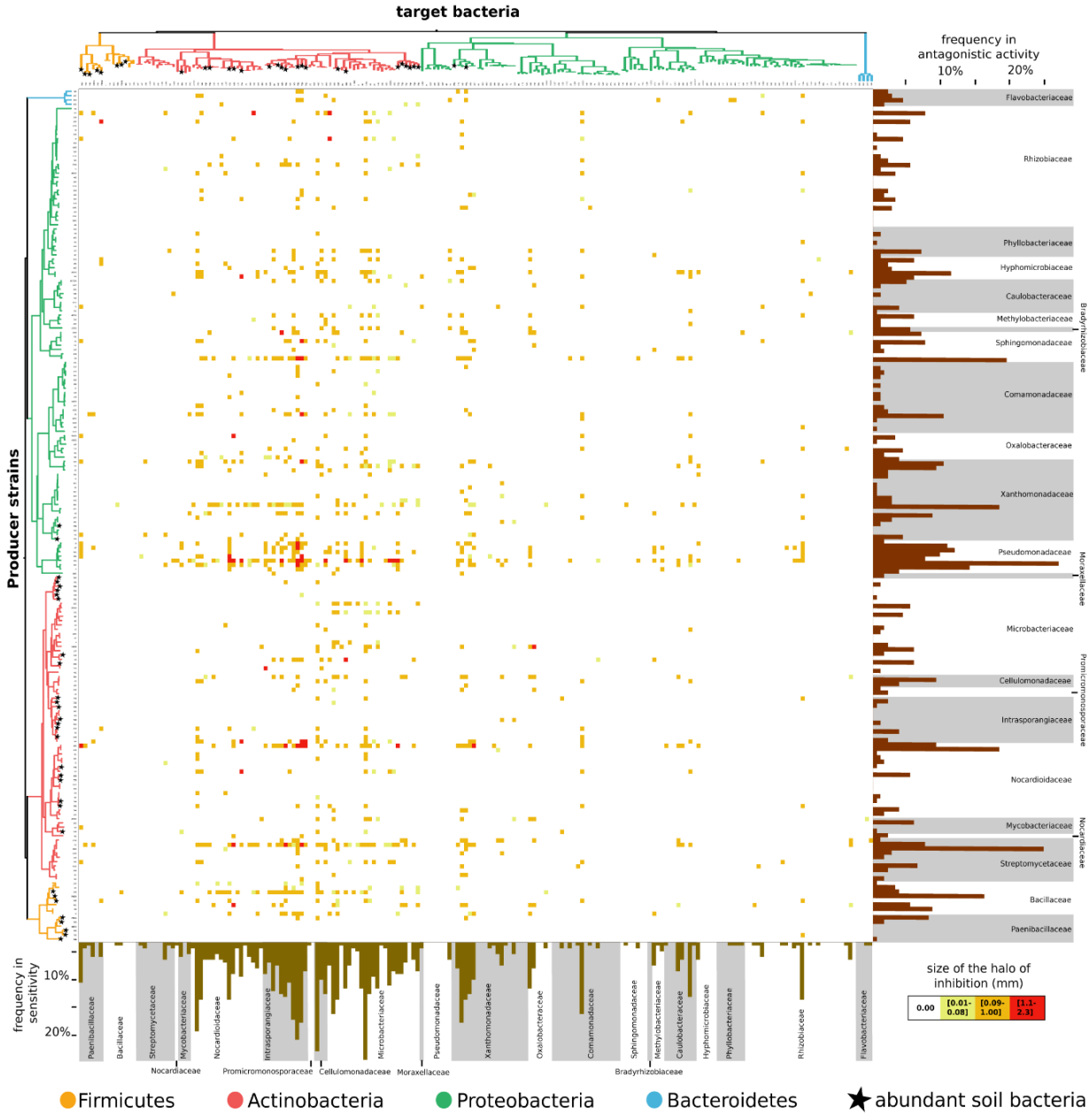


Figure 9| **The bacterial isolates show different inhibition frequencies.**

The graph shows 39,204 tested interactions and indicates the 1,011 mutual inhibitions. Bacteria that produce antimicrobials are indicated in the y-axis as producer strains and bacteria that are sensitive to secreted antimicrobial are plotted in the x-axis as target bacteria. Both trees correspond to 16s rRNA gene phylogenetic distance between the isolates. Stars in the phylogenetic trees refer to abundant soil bacteria. Saturation in the red color indicates the size of the halo of inhibition. Several bacteria show a level of antagonistic activity or sensitivity greater than 15%.

Several Actinobacteria community members are out-competed. Although inter-bacterial inhibitions are uncommon, root-derived and abundant soil bacteria engage in antagonistic interactions through the secretion of antimicrobials and 1,011 inhibitions are reported among community members. In order to reveal the competitive potential of *Actinobacteria*, *Bacteroidetes*, *Firmicutes* and *Proteobacteria,* I first aggregated the inhibitions scores at phylum level. Notably, ~70% of observed inhibitions target *Actinobacteria* whereas only ~28% affect *Proteobacteria* (**Figure 10**, **panel-A**). In contrast, *Proteobacteria* explains more than 60% of overall observed antagonistic activity compared to *Actinobacteria* that explains only ~29% (**Figure 10**, **panel-A**). Taken together, these data show striking differences in the antagonistic profiles between *Proteobacteria* and *Actinobacteria* and indicate that *Actinobacteria* is more sensitive to secreted antimicrobials. It is important to highlight that *Proteobacteria* is found reproducibly across several plant species and significantly enriched and dominant in the roots according to culture-independent profiling (Dombrowski *et al.,* 2017, Edwards *et al.,* 2015, Peiffer *et al.,* 2013, Bulgarelli *et al.,* 2012). It is then plausible to hypothesize that antagonistic interactions among bacteria may shape microbial assemblages at the root/soil interface without excluding the hypothesis that isolates from *Proteobacteria* are better adapted to that particular niche.

Since the number of isolates is not even across the four phyla, I normalized the inhibition scores by total number of interactions within each phylum (**Figure 10**, **panel-B**). The bar-plots portrayed in **Figure 10**, **panel-B** show the remarkable higher level of sensitive of *Actinobacteria* compared to all other bacterial phyla. Nonetheless, frequencies in antagonistic activity are relatively close to one another across the four phyla and vary between two to three percent. Our data suggest that *Actinobacteria* is disproportionately sensitive to secreted antimicrobials, whilst more stable frequencies of antagonistic activity are detected across all four phyla. To further investigate the sensitivity of *Actinobacteria* and reveal intra- and inter-phylum inhibitions, I scored the inhibition by only the product of interactions of a phylum to the corresponding interacting phylum (*i.e. Actinobacteria versus Actinobacteria*, *Actinobacteria versus Firmicutes*, *etc.*). The network in intra- inter-phylum inhibitions is illustrated in **Figure 10**, **panel-C** and in the following *Actinobacteria* is referred by A, *Bacteroidetes* by B, *Firmicutes* by F and *Proteobacteria* by P for simplification. The analysis of the network depicted **Figure 10**, **panel-C** indicates that all phyla engage in intra- and inter-phylum inhibition, except P do not inhibit B, no cross inhibitions between F and B and no intra-phylum inhibition is observed in B. Interestingly, the network indicates a strong inter-phylum inhibition directed against *Actinobacteria* by F (12.1%), P (5.3%) and B (3.8%) (**Figure 10**, **panel-C**). In contrast *Actinobacteria* do not show high antagonistic activities against bacteria from other phyla (percentages of inhibition range from 0.7 to 1.4%) (**Figure 10**, **panel-C**). Inspection of intra-phylum inhibition reveals extremely low frequencies for F (1%) and P(1.4%), whereas a strikingly high level of antagonism is observed within *Actinobacteria* members (4%), which indicate extensive competition among *Actinobacteria* members (**Figure 10**, **panel-C**). The analysis of intra- and inter-phylum antagonistic interactions indicates that *Actinobacteria* are more sensitive to secreted antimicrobials by other bacterial phyla as by *Actinobacteria* community members

(**Figure 10**, **panel-C**). Importantly, inter-phylum analysis reveals that antagonistic activity of *Actinobacteria* hardly exceed 1% while its inter-phylum sensitivity is never below 3.8% (**Figure 10**, **panel-C**). Most strikingly, *Actinobacteria* shows by far the highest score of intra-phylum inhibition among the two bacterial phyla that show intra-phylum inhibition (**Figure 10**, **panel-C**). It is plausible that several *Actinobacteria* isolates are out-competed by other community members during host-associated microbiota establishment.

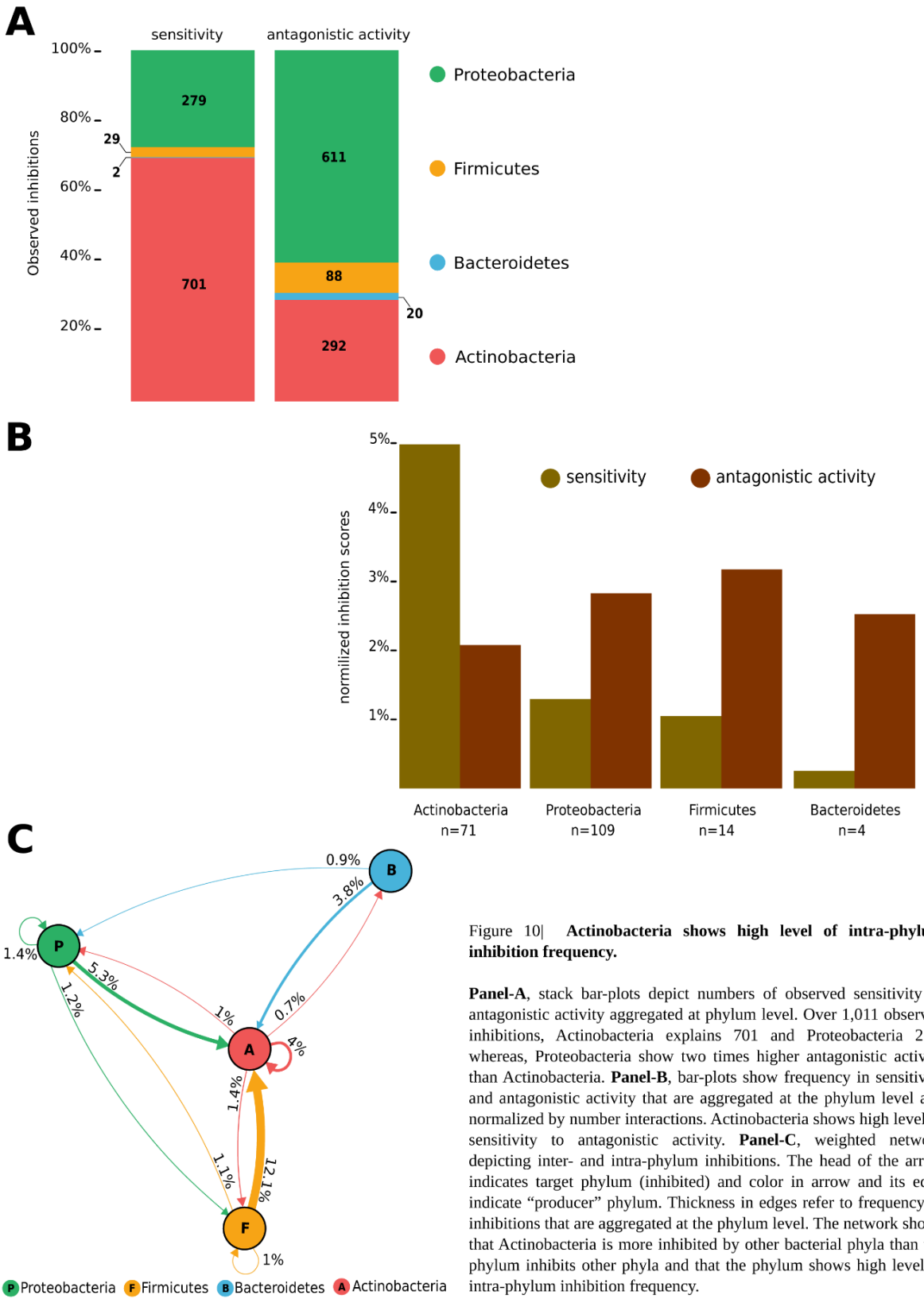


Figure 10| **Actinobacteria shows high level of intra-phylum inhibition frequency.**

Panel-A, stack bar-plots depict numbers of observed sensitivity or antagonistic activity aggregated at phylum level. Over 1,011 observed inhibitions, Actinobacteria explains 701 and Proteobacteria 279, whereas, Proteobacteria show two times higher antagonistic activity than Actinobacteria. **Panel-B**, bar-plots show frequency in sensitivity and antagonistic activity that are aggregated at the phylum level and normalized by number interactions. Actinobacteria shows high level of sensitivity to antagonistic activity. **Panel-C**, weighted network depicting inter- and intra-phylum inhibitions. The head of the arrow indicates target phylum (inhibited) and color in arrow and its edge indicate "producer" phylum. Thickness in edges refer to frequency of inhibitions that are aggregated at the phylum level. The network shows that Actinobacteria is more inhibited by other bacterial phyla than the phylum inhibits other phyla and that the phylum shows high level of intra-phylum inhibition frequency.

Competitive interactions among and between distantly and closely related community members. The analysis of inter-strains antagonistic interactions indicated that *Actinobacteria* is a highly sensitive to secreted antimicrobials bacterial phylum and displays a high level of intra-phylum antagonistic interactions. Intra-phylum antagonistic interactions can result from both intra- and inter-family antagonistic interactions. Moreover, since several bacterial isolates have been predicted to harbor BGCs involved in the synthesis of bacteriocin or microcin, known antimicrobials that mediates often the inhibition of closely related species (Riley and Gordon, 1999), it is very likely to observe antagonistic interactions between closely related community members. In order to explore intra- and inter-family inhibitions, I first aggregated the inhibition scores at the family level and normalized each sum by the number of inter- or intra-family interactions. The heatmap in **Figure 11** displays intra- and inter-family inhibition percentages that have been reported in the ABBA screen. This analysis highlights four families in *Actinobacteria* that are sensitive to antimicrobials secreted by broadly diverse strains. In contrast, only one *Proteobacteria* family, *Xanthomonadaceae,* that shows similar broad sensitivity (**Figure 11)**. Interestingly, most of high inhibition percentages (>20%) are mainly reported in *Actinobacteria* and correspond almost in all cases to inter-family interactions (**Figure 11**). Although inter-family inhibitions seem to occur more often than intra-family inhibitions, 12 bacterial families show intra-family antagonistic interactions with ranging percentages from 0.9 to 22% (**Figure 11**). Interestingly, three out five intra-family inhibitions observed in *Actinobacteria* show a percentage ≥11% (*Mycobacteriaceae*, *Cellulomonadaceae*, *Microbacteriaceae),* whereas all remaining intra-family inhibitions remain ≤5% (**Figure 11**). To further corroborate these data, I analyzed the phylogenetic distances (based on full-length 16S rRNA gene) between producer and target bacteria (**Figure 12 panel-A**) and *vice versa* (**Figure 12 panel-B**). This analysis identifies the inhibitions, whether antagonistic activity or sensitivity, spectra of the isolates. Using this strategy, it is possible to reveal limited or broad antagonistic activity that each isolate exert on other microbiota members in a phylogenetic context. Since the x-axis shows phylogenetic distances, bacteria that share more than 87% of sequence similarity are more likely to be from a same family, bacteria that share more than 97% of sequence similarity are more likely to be the a same species. Although only a few inhibitions are directed toward closely related members, it is clear from the **Figure 12 panel-A** that *Proteobacteria* isolates tend to inhibit more distantly related isolates than closely related strains. Interestingly, several isolates belonging to different *Proteobacteria* families appear to be sensitive to *Bacteroidetes* isolates. In turn, isolates belonging to the family *Flavobacteriaceae* (*Bacteroidetes*) appear to be highly resistant since they are inhibited by only two *Actinobacteria* isolates (**Figure-12, panel-B**). Most strikingly, both panels **A and B** in the **Figure 12** demonstrate the amplitude of antagonistic interactions that occur between *Actinobacteria* isolates that share more than 87% sequence similarity or with even more related strains (**Figure 12, panels A and B**). The inhibition of relatives is mainly, but not exclusively, observed in *Microbacteriaceae*, *Cellulomonadaceae*, *Intrasporangiaceae* and *Nocardioidaceae* (**Figure 12, panels**

A and B). Although few inhibitions of closely related members are detected in *Xanthomonadaceae* and *Rhizobiaceae*, *Actinobacteria* isolates belonging to the four above-mentioned families are by far the most affected by antimicrobials produced by closely relative cells (**Figure 12, panels A and B**). Taken together, these data reveal that a majority of *Actinobacteria* isolates are not only sensitive to antimicrobials secreted by phylogenetically distant bacteria but also show higher intra-family inhibition. Furthermore, these data suggest that *Actinobacteria* isolates engage in a fierce competitive interactions with congeners during community establishment.

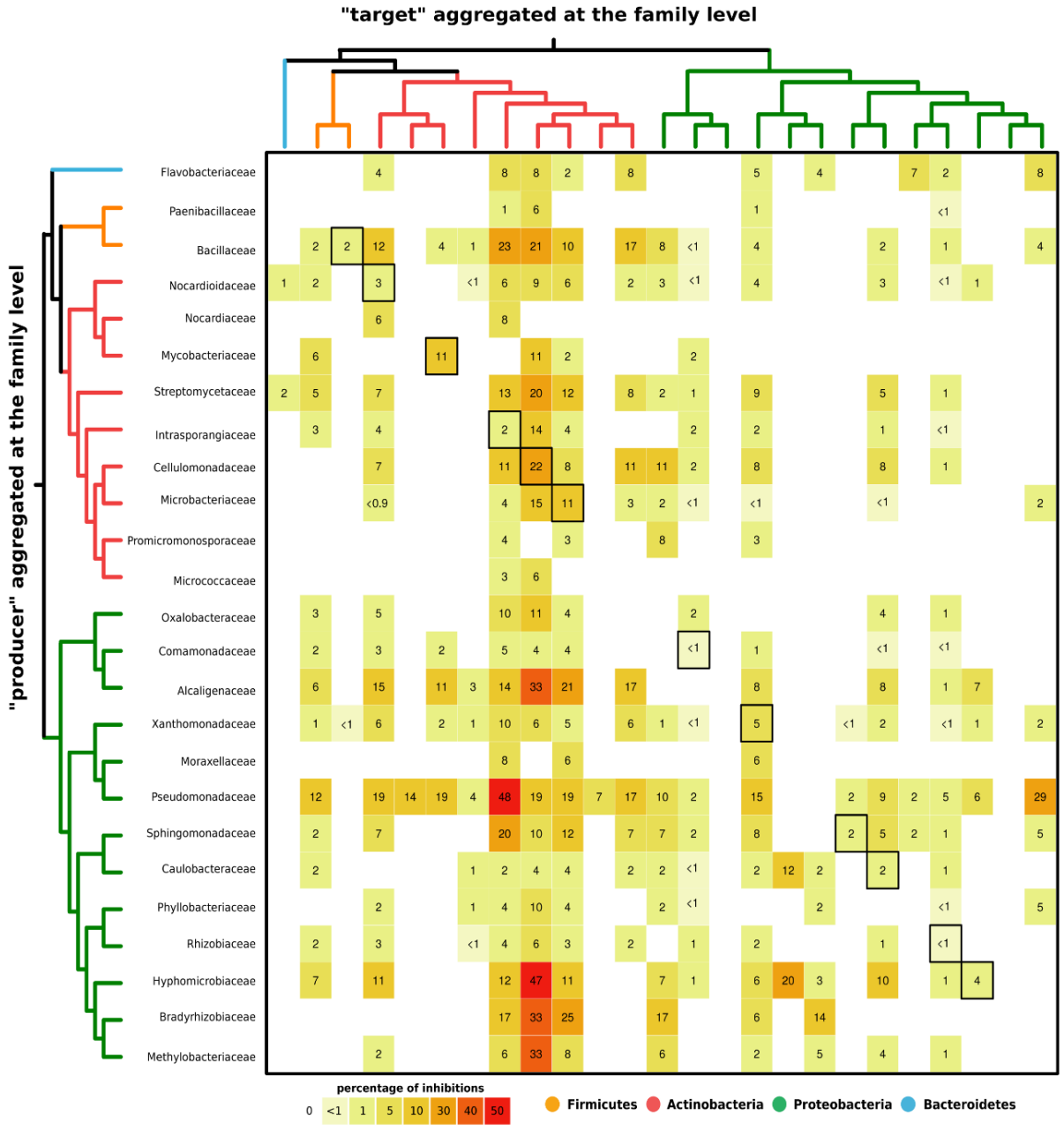


Figure 11| **Network of inter- and intra-family antagonistic interactions.**

The heat-map depicts inter- and intra-family network of inhibitions. Intensity in the red gradianet indicates increase in the frequency of inhibition. "Producer" bacterial families are indicated in y-axis and "target" bacterial families are indicated x-axis. Color-code in phylogenetic trees indicate bacterial phyla. Four families in Actinobacteria, *Nocardioidaceae*, *Intrasporangiaceae*, *Cellulomonadaceae* and *Microbacteriaceae* and one Proteobacteria family, *Xanthomonadacea*, are inhibited by a wide rage of bacterial families. Several bacterial families show intra-family inhibitions, however this inhibitions are more frequent in Actinobacteria.

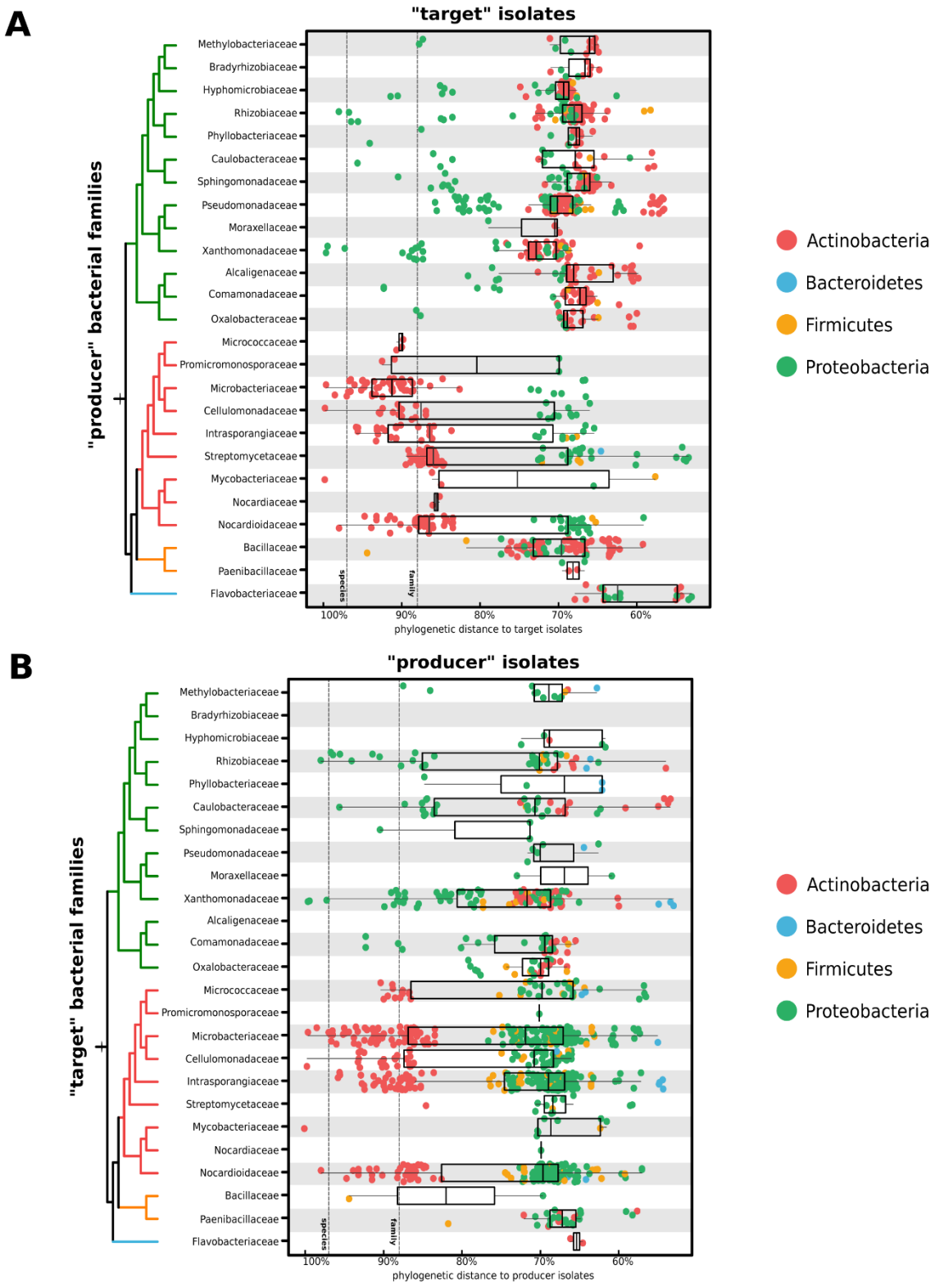


Figure 12| **Several Actinobacteria isolates inhibit closely-related species.**

Panel-A, the plot shows phylogenetic distance of bacteria that are inhibited by the bacterial isolates indicated in y-axis and grouped by family. Each dot in the graph corresponds to one sensitive isolate. **Panel-B**, the plot shows the phylogenetic distance of target isolates grouped at family and indicated in y-axis to their antagonizers. Each dot in the graph corresponds to one producer isolate. This analysis indicates that several Actinobacteria isolates inhibit closely-related, below family level, isolates.

Correlation network of antagonistic profiles. Our ABBA screen revealed that more than half of the bacterial isolates inhibit or are inhibited by at least one isolate and that numerous isolates show multiple inhibitions. Therefore, it is plausible that bacteria could share a certain degree of similarity in their antagonistic activity or sensitivity profiles. To address this question, I calculated correlation among sensitivity profiles and among antagonistic activity profiles for all the isolates that show at least one sensitivity or antagonistic activity, respectively. Network plots depicted in **Figure 13** show correlations in sensitivity and antagonistic activity profiles. Both networks show more correlations under the cutoff 0.95 than equal to one (**Figure 13, panels A and B**). From the sensitivity correlation network, two cloud of nodes can be distinguished; the first corresponds to a group of nodes that are more connected and plotted at the left periphery, a second cluster plotted at the center of the network and over-dominated by weak connections (**Figure 13, panel-A**). Within the former cloud, several small clusters of isolates show a high positive coefficient of correlation (r =1), however these isolates are inhibited by only one isolate (**Figure 13, panel-A**). Regarding the latter cluster, all nodes representing the bacterial isolates are inhibited by more than one bacterium (**Figure 13, panel-A**). Although there are much less correlations between the isolates within the later cluster, it is interesting to highlight the strong similarity in the sensitivity profile observed between the isolates n°198D2, 122 and 112D2 or between the *Rhizobium* n°483D2 and the *Methylobacterium* n°483D1 (**Figure 13, panel-A**). The analysis of the sensitivity correlation network indicates that most of the isolates show rather different patterns in their sensitivity profiles and only exceptionally the bacteria have exactly similar profiles in sensitivity when they are inhibited by more than one strain. Regarding the antagonistic activity correlation network, a similar topology as the sensitive correlation network has been observed: 1- a first cluster, located at the top periphery of the network, is constituted of nodes that show a strong positive correlation between bacteria that have a single and identical inhibitory activity profile, 2- a second cluster of nodes plotted at the center of the network and represented by isolates that inhibit more than one bacterium (**Figure 13, panel-B**). Although most of the nodes are weakly connected in the latter cluster, few strong positive correlations can nonetheless be observed (**Figure 13, panel-B**). These correlations indicate that considered bacterial isolates show similar profiles in their inhibition patterns. More interestingly, correlations in antagonistic activity do not systematically occur between bacteria from a same phylum, as illustrated in **Figure 13, panel-B.** For instance, the antagonistic profile of the *Proteobacteria* isolate n°217 strongly correlates with the *Actinobacteria* isolate n°485 (**Figure 13, panel-B**). Very few positive correlations in antagonistic activity were identified between closely related bacteria, (*i.e.* family and species level), indicating that phylogenetic distance is not a good predictor of the killing spectrum of individual bacterial isolate (**Figure 13, panel-B**). The fact that very few bacteria have similar patterns of antagonism suggests a large a genetic variation in competitive mechanisms across bacterial families and even species/isolates. This diversity of antagonistic profiles is likely important for stabilizing microbial networks since both species richness and inhibition pattern diversity may maintain ecological communities. However, a large percentage of the isolates show rather weak connections. Both sensitivity

and antagonistic activity correlation networks considered, bacteria tend to have different or weakly overlapping sensitivity and antagonistic activity profiles. Nonetheless, few examples do not follow the rule of thumb and only few isolates that inhibit more than one bacterium show a strong correlation in their inhibition profiles.

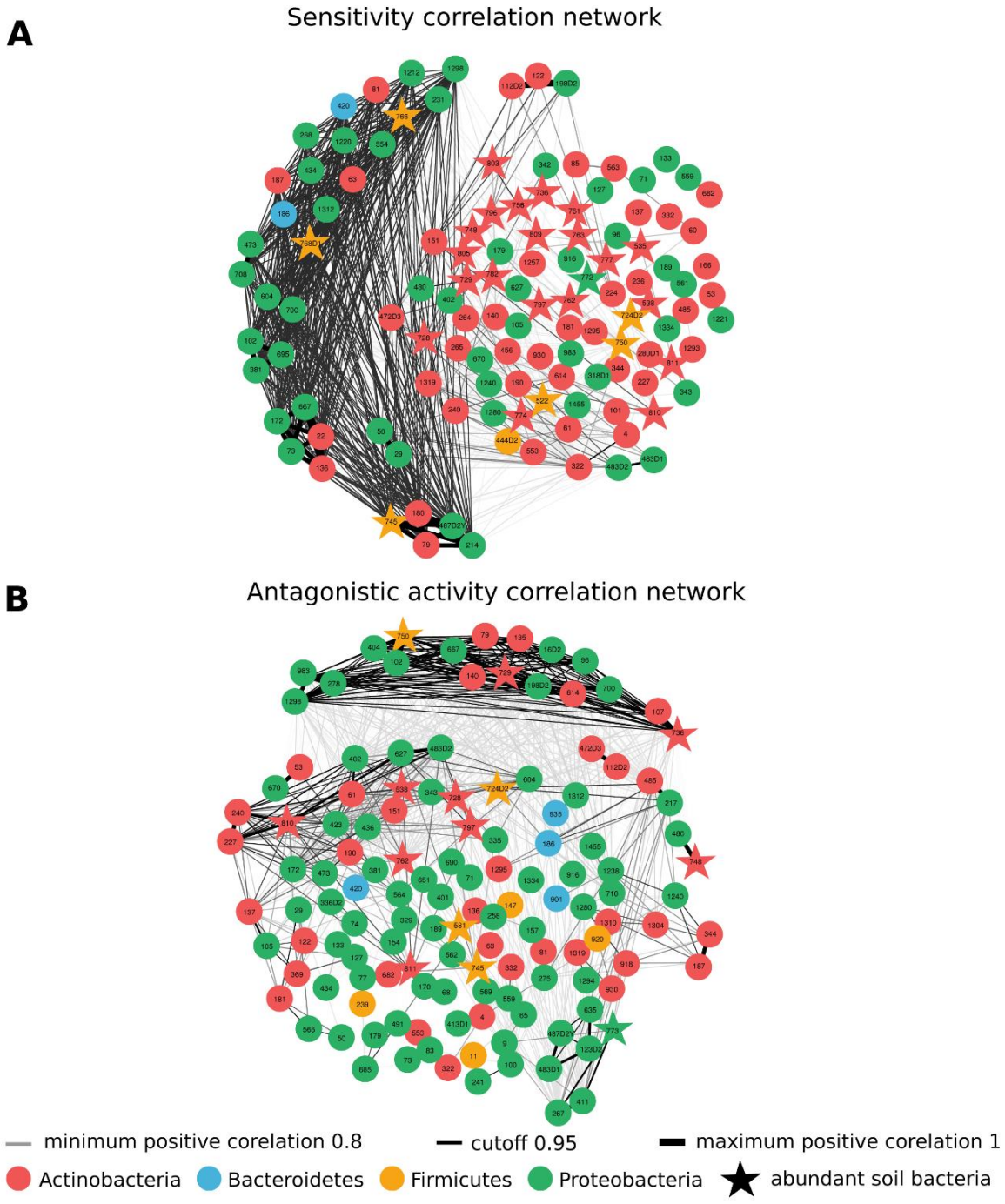


Figure 13| **Correlation network in the inhibition profiles of bacteria**

Panel-A and **-B** depict respectively sensitivity and antagonistic activity correlation network. Nodes represent bacterial isolates. Color-code indicates phylum. Abundant soil bacteria are referred by star in both panels. Edges show positive tetrachoric correlation coefficient in sensitivity, **panel-A**,or in antagonistic activity profile, **panel-B**, computed on binary dataset (presence/absence of inhibition). All correlation coefficients are above 0.8 with FDR-corrected significance level of 0.05. Cutoff in correlation coefficients is 0.95. Edges color, from bright to dark, and thickness, from thin to large, correspond to degrees from low to high in correlation coefficient. Graphical LASSO method was used to estimate the network structure. All represented networks show few nodes with strong correlation coefficient (~1) and several nodes with intermediate correlation coefficient (< 0.95).

Validation of the ABBA screen To evaluate the reproducibility of our ABBA screen, a random set of bacterial isolates were chosen and re-screened for antagonistic interaction. Since several bacterial isolates are spotted on top of a lawn bacterium, the adjacencies of these isolates are also randomized to ovoid bacteria being spotted exactly in the same position as in the original screen. The validation screen includes several bacterial isolates belonging to the four phyla, including root-derived and abundant soil bacteria. Although several bacteria could not grow throughout the validation screen, 90 isolates have been successfully tested as spot bacteria and 83 as lawn bacteria (**Figure 14**). Over 7,470 validation tests, more than 7,000 of the interactions are successfully reproducible and 352 interactions are not reproducible (**Figure 14**). The fraction of irreproducibility is ~5% and mainly explained by 60% of new detected inhibitions and 40% of missing inhibitions (**Figure 14**). The new detected inhibitions are interesting observations since these new inhibitions could be explained by new isolates' adjacency as already demonstrated by Abrudan *et al.*, 2015 as well as by technical irreproducibility. The validation screen allows to evaluate the reproducibility of the ABBA screen and could confirm that 95% of the interactions are reproducible

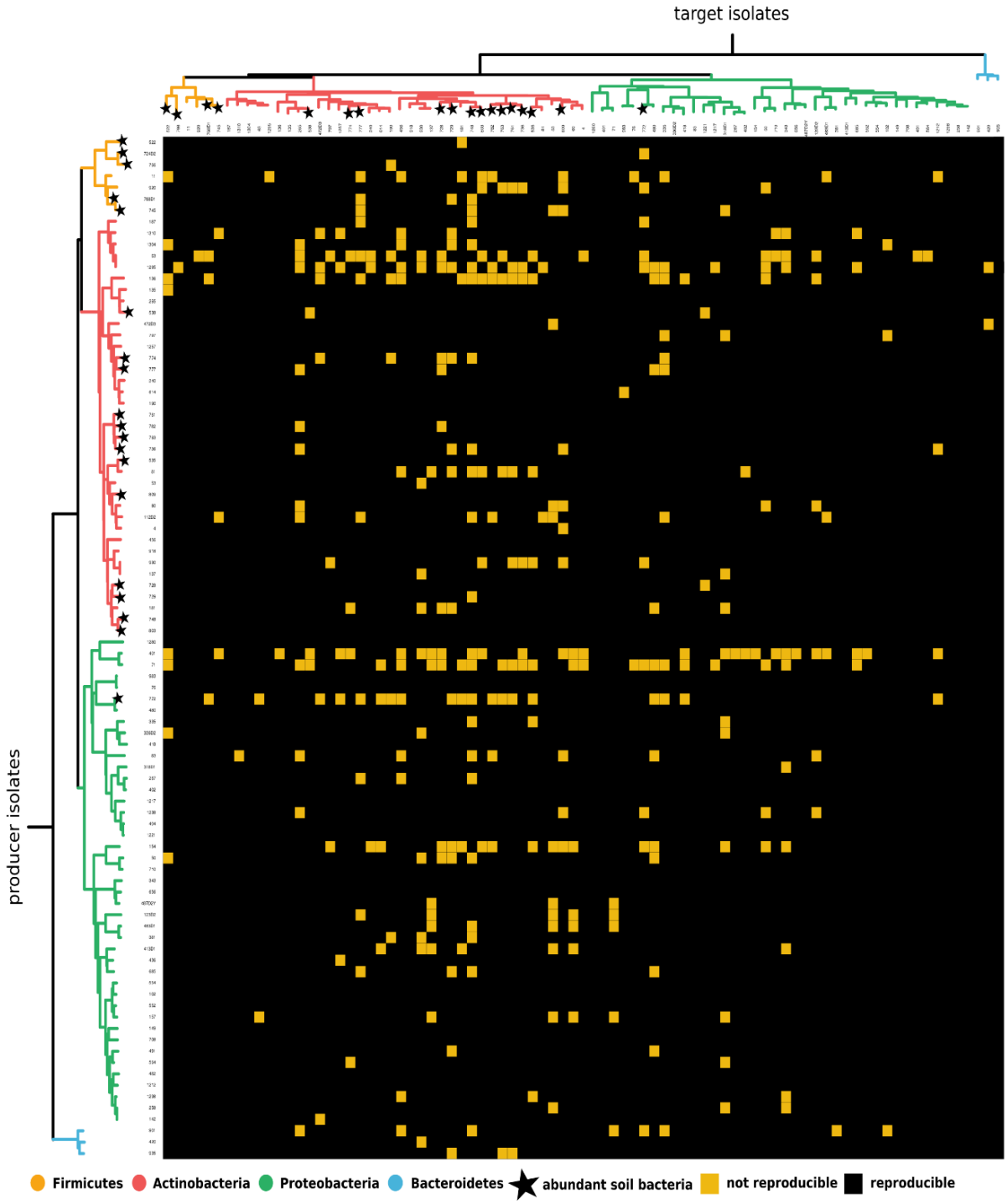


Figure 14| **Validation of the antagonistic bacteria-bacteria interactions assay**

A random subset of bacterial isolates have been taken to re-screen for mutual inhibitions. Spot bacteria that are tested for antagonistic activity are indicated in the y-axis and the color in the tree refers to the phylum. Target bacteria are indicated in the x-axis on the top. Abundant soil bacteria are indicated by a star. 95% of the interactions indicated in the graph by a black cell are reproducible and ~5% of the interactions were not reproducible.

Intra- and inter-specific antagonistic interactions. The bacterial isolates used in this study (*i.e.* genome mining for BGCs, metabolome analysis and screen for antagonistic activity) belong to 2 ecologically-delineated bacterial groups; root-associated bacteria represented by 167 isolates that have been reproducibly found by high-throughput sequencing methods associated with the roots of *A. thaliana* (Bai *et al.*,2015). These bacteria are designated as root-derived bacteria. The remaining isolates used in this study have been found to be enriched in the bulk soil and hardly detected in the roots of *A. thaliana* (Bai *et al.*,2015), and they are designated by abundant soil bacteria. Both former and latter bacterial groups have been shown to engage in antagonistic interactions. Interestingly, it has been shown that antagonistic interactions can be more frequent between bacteria that belong to different micro-habitat "inter-specific competition" than between bacteria that belong to a same micro-habitat "intra-specific competition" (Cordero *et al.*,2012). Although abundant soil bacteria are under-represented compared to the root-derived bacteria, we hypothesize that inter-specific antagonistic interactions (i.e. between soil and root isolates) are more frequent than intra-specific interactions (within soil isolates or within root isolates). Furthermore, since plant roots are densely colonized by bacteria and abundant soil bacteria are almost excluded from the rhizoplane (roots surface), I tested the hypothesis that abundant soil bacteria are highly sensitive to secreted antimicrobials by root-derived bacteria. In order to study intra- and inter-specific antagonistic interactions, I re-analyzed the data and aggregated the inhibition scrores according to root-derived and abundant soil bacteria groups. The heatmap plot in **Figure 15, panel-A** shows antagonistic interactions between and within root-derived and abundant soil bacteria. The analysis of intra- and inter-specific antagonistic interactions reveals that bacterial isolates with a frequency of antagonistic activity more than 5% are almost exclusively root-derived bacteria. Moreover, the aggregated antagonistic activity index of the root-derived bacteria is two times higher than the abundant soil bacteria index. Thus, the root-derived bacteria tend to inhibit more isolates than abundant soil bacteria. Nonetheless, it is still yet possible that the overall sensitivity are comparable between root-derived and abundant soil bacteria. In order to further investigate the sensitivity between root-derived and abundant soil bacteria, I have aggregated the sensitivity of both bacterial groups. Although, root-derived bacterial isolates show high frequency in sensitivity, abundant soil bacteria are overall two times more sensitive to antagonistic activity than root-derived bacteria. Moreover, the analysis of inhibitions between and within former and latter bacterial groups indicates that root-derived bacteria inhibit more frequently abundant soil bacteria and most strikingly that abundant soil bacteria show more intra-specific inhibitions than inter-specific. Taken together, these data indicate that root-derived bacteria are more resistant to secreted antimicrobials and abundant soil bacteria are more sensitive to secreted antimicrobials. Since the sampling in abundant soil bacteria is low compared to root-derived bacteria, it is important to further sample taxonomically diverse abundant soil bacteria in order to confirm this observation.

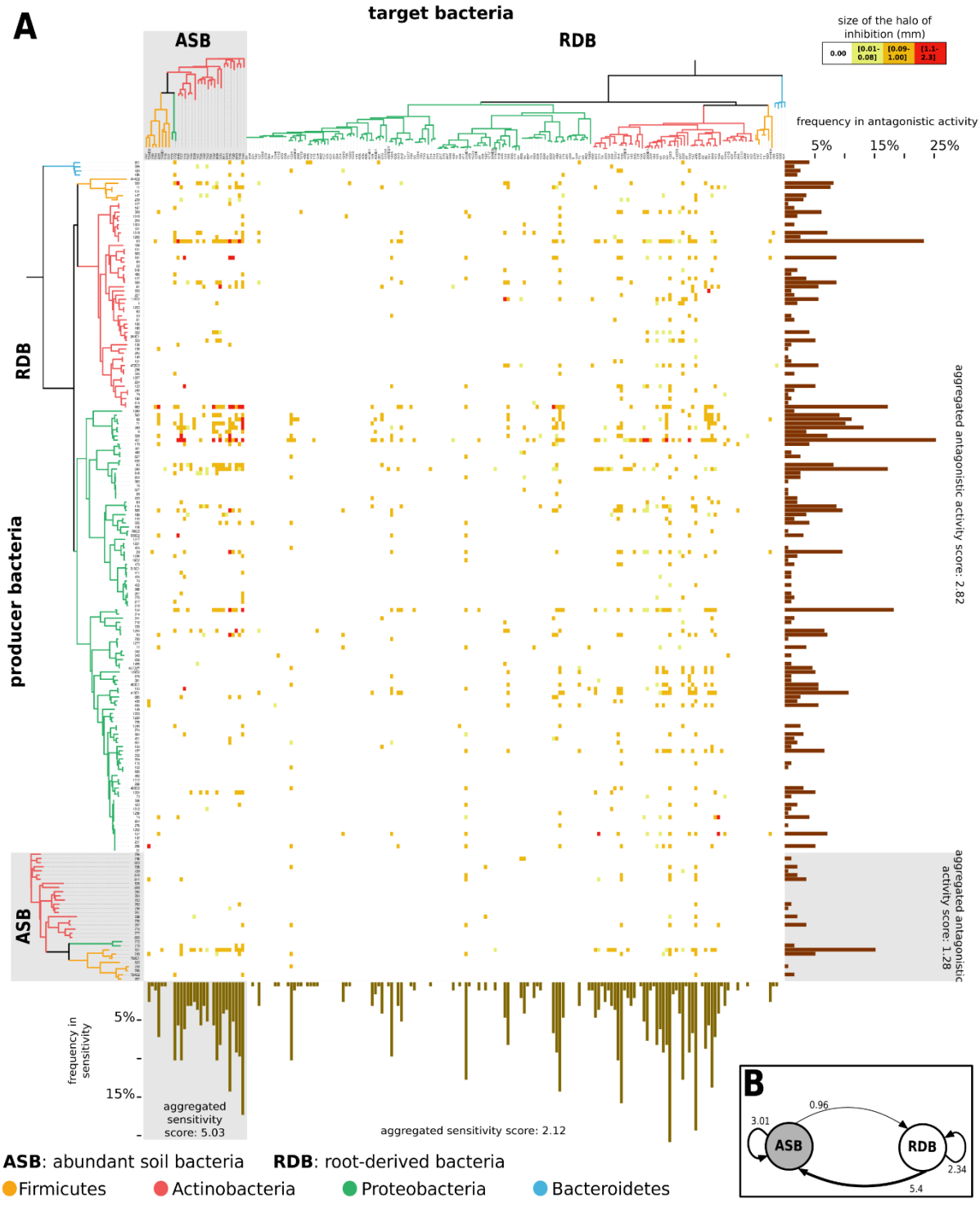


Figure 15| **Abundant soil bacteria are out-competed.**

Panel-A, The heatmap depicts the observed inhibitions within and between root-derived and abundant soil bacteria. The aggregated frequency in sensitivity of abundant soil bacteria is two times higher than the aggregated frequency of root-derived bacteria. But, the aggregated frequency of antagonistic activity of root-derived bacteria is higher than the aggregated frequency in antagonistic activity of abundant soil bacteria. **Panel-B**, the two components network shows the frequency of inhibitions within and between abundant soil bacteria and root-derived bacteria. Intra-specific inhibition frequency of the abundant soil bacteria in more important than the inter-specific inhibition frequency. Abundant soil bacteria are highly sensitive to bacterial antagonistic activity.

Defining two groups of strains with a contrasting competiveness potential. Bacteria live in dense microbial communities where competitive interactions do not only affect their survival but also are thought to alter the community diversity and structure (Maida *et al.*, 2015, Czárán *et al.*, 2002). Revealing competitive potential of individual microbiota members is needed to identify highly competitive members and test their role in altering the community structure and diversity under different experimental systems. The ultimate goal of our multifaceted approach (*i.e.* mining genomes for BGCs, analysis of secreted metabolites and screen for antagonistic activity) is to uncover the competitive potential of each bacterium from a comprehensive culture collection and to identify highly competitive and highly sensitive community members. By defining highly competitive and highly sensitive bacteria, it becomes possible 1- to disentangle the influence of each contrasted group in the establishment of microbial communities *in vitro* and *in planta* and 2- to test whether competitiveness can alter community diversity and structure. In order to define highly competitive bacteria, I plotted the sum of observed antagonistic activity for each isolate by its antagonistic activity degree (*i.e.* average size in the measured halo of inhibitions for each producer strain) (**Figure 16 panel-A**), and I plotted the sum of observed sensitivity to secreted antimicrobials for each isolates by its sensitivity degree (*i.e.* average size in the halos of inhibitions for each targer strain) (**Figure 16 panel-B**) to define highly sensitive bacteria. In order to select top competitive and top sensitive strains, I fixed two thresholds within each group and selected members that are above both fixed thresholds. Thresholds of antagonistic activity or sensitivity are set to 2.5 times the median value in antagonistic activity or sensitivity, respectively. Similarly, the threshold in the degree in antagonistic activity and sensitivity are fixed as 2.5 times the mean value of each phenotype (antagonistic activity or sensitivity), respectively. Additionally, strains that are selected as highly competitive should not belong to highly sensitive strains and *vice versa*. By analyzing the network of inhibitions, two groups of bacterial strains with contrasting competitiveness potential are defined: 1- 13 highly competitive community members that show high antagonistic activity and very low sensitivity to antagonistic activity, 2- 13 highly sensitive community members that are highly sensitive to secreted antimicrobial and very limited in their antagonistic activity. Using the above-mentioned threshold, I defined two non-overlapping groups, each composed of 13 bacterial strains with either highly competitive or sensitive potential (**Figure 16, panel-A and -B)**. Interestingly, all selected highly competitive members are root-derived bacteria and a majority of them belong to *Proteobacteria* (**Figure 16, panel-A),** whereas highly sensitive members are mostly abundant soil bacteria and predominantly *Actinobacteria* (**Figure 16, panel-B**).

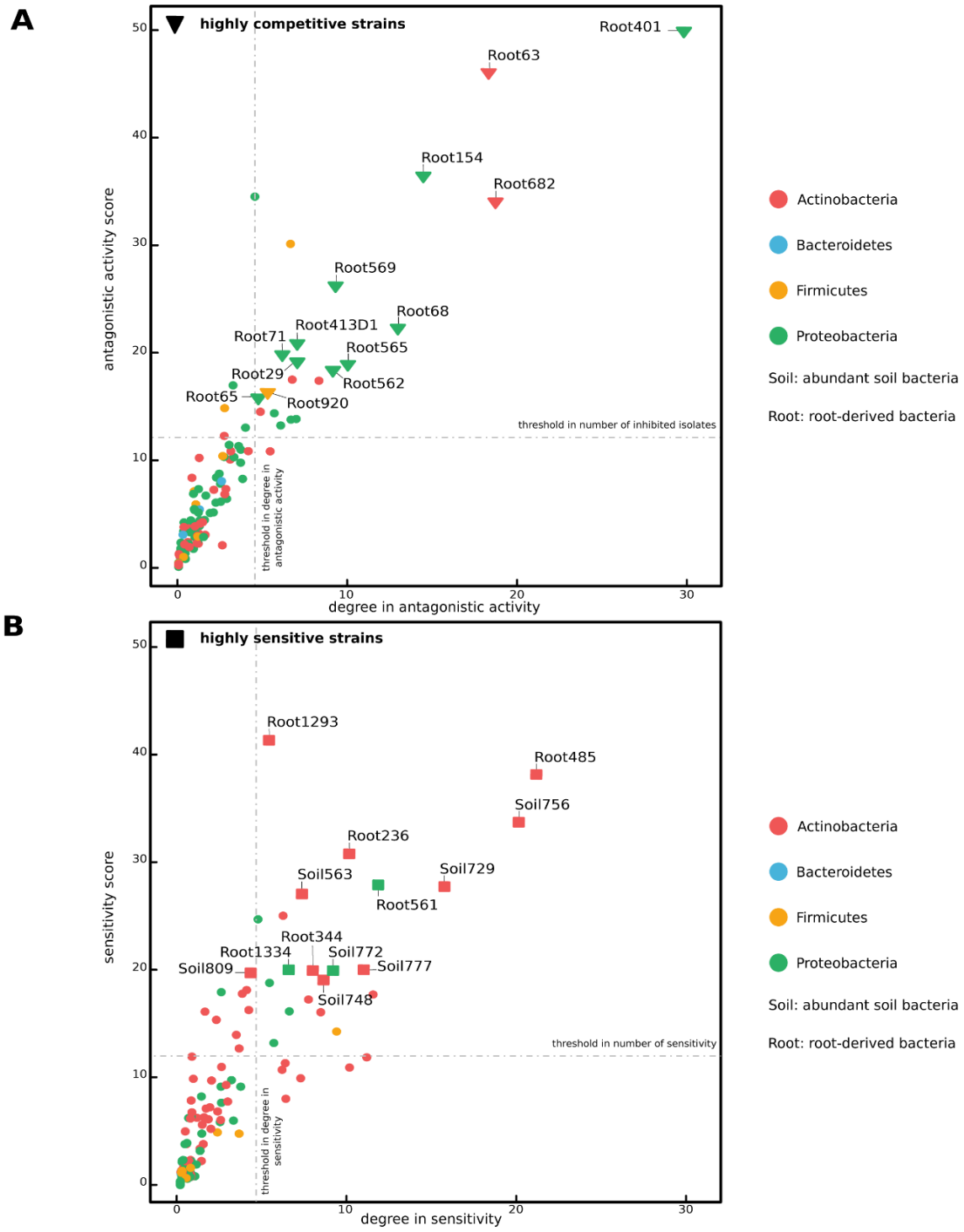


Figure 16| **Defining two groups of bacterial strains with contrasting competitive potential**

Panel-A, defining a group of isolates with high competitive potential. The scatter plot is obtained by projecting number of observed antagonistic activity for each isolate by its degree in antagonistic activity (mean value of measured size in the halos of inhibitions). Highly competitive bacteria are indicated by inverted triangle. **Panel-B**, defining a group of isolates with low competitive potential. The scatter plot is obtained by obtained by projecting number of observed sensitivity for each isolate by its degree in sensitivity (mean value of measured size in the halos of inhibitions). Highly sensitive bacteria are indicated by square shape.

In order to cross-link between the genomes analysis and the screen for antagonistic interactions, I plotted antagonistic activity and sensitivity scores and BGCs numbers for each bacterium in a phylogenetic tree. The raw data are indicated in **Supplementary Table 3**. From this analysis, it is interesting to highlight that only one isolate from the highly competitive bacteria show a total BGCs count below two, whereas no less than six highly sensitive bacteria have less than three biosynthetic clusters (**Figure 17, panel-A**). Moreover, bacteria with more than seven BGCs show hardly high level of sensitivity (**Figure 17, panel-A**). To further reveal whether isolates with high BGCs content are more likely to be more competitive than bacteria with low BGCs count and that bacteria with less BGCs are more likely to be sensitive to antagonistic activity, I plotted the probability to observe antagonistic activity or sensitivity, $P(I|n)$, according to predicated number of BGCs (**Figure 17, panel-B**). The analysis of the probability of inhibitions indicates that bacteria with less BGCs tend to be more sensitive and show low antagonistic activity. Interestingly enough, the sensitivity drops quickly by the increase number of predicted BGCs. Although antagonistic activity increase in isolates with more BGCs, it seems to reach a plateau indicating that high frequency in antagonistic activity do not increase linearly with the number of BGCs (**Figure 17, panel-B**). As for unique nodes, highly competitive bacteria have in total five times more unique nodes than highly sensitive bacteria (**Figure 17, panel-A**). Taken together, our data indicate that bacterial isolates with low BGCs count are more likely to be sensitive to secreted antimicrobial by other community members then bacteria with high BGCs count. It is plausible that high BGCs number and diversity could confer to a bacterium resistance and/or protection against antimicrobials on top of the competitive potential. It is however, interesting to highlight that highly competitive bacteria in this study tend to have more genetic and chemical features than highly sensitive members.

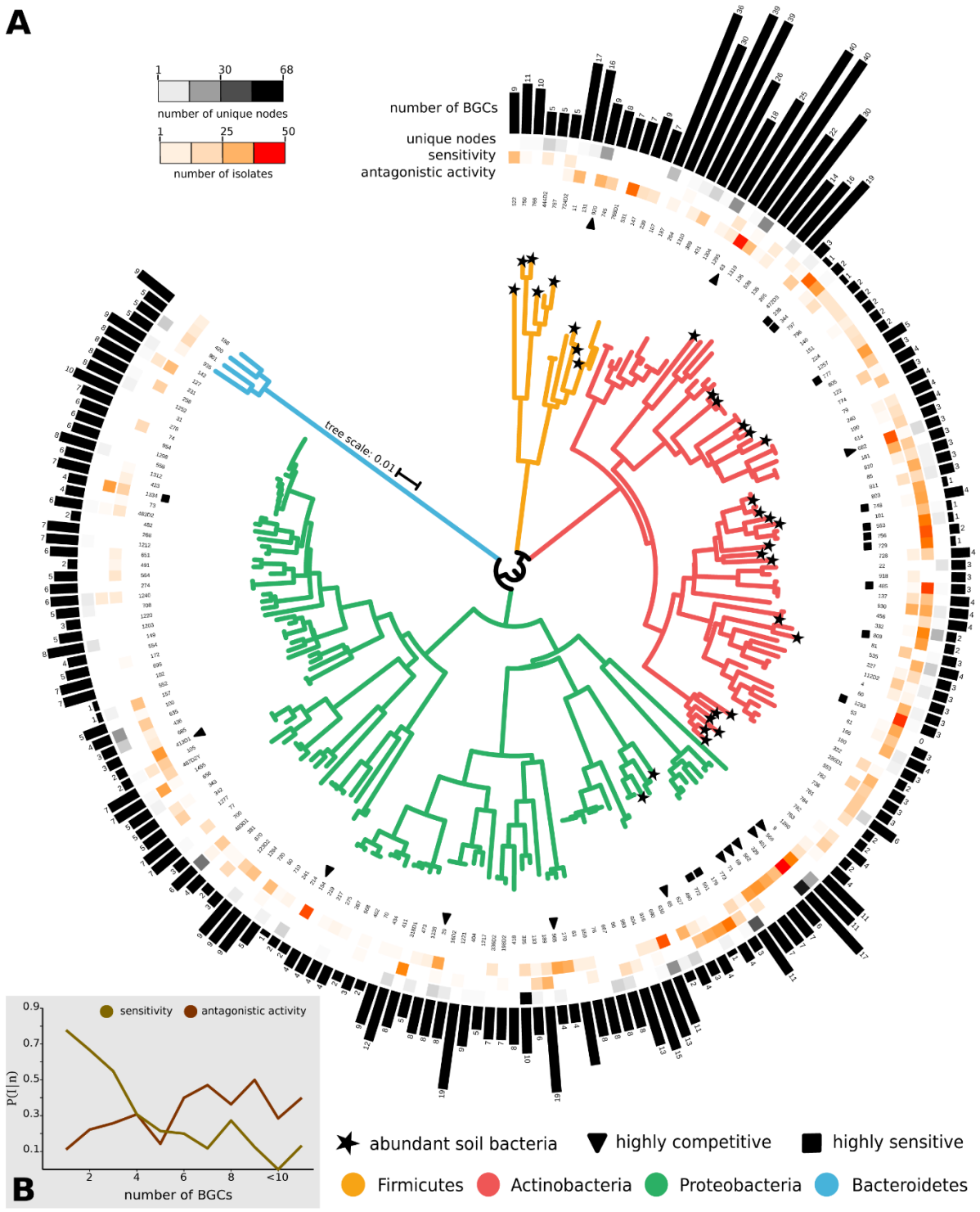


Figure 17| **Overview of genome, metabolome and "antagonome" studies.**
Panel-A, the circular phylogenetic tree shows bacterial isolates included in genome, metabolome and screen for mutual inhibitions studies. Colors in tree branches' indicate phyla and stars indicate abundant soil bacteria. Highly competitive and highly sensitive bacteria are indicated in the tree respectively by inverted triangle and square. Antagonistic activity and sensitivity scores of each isolate is indicated by the red gradient color. Gradient from gray to black shows number of unique chemical features for each isolates. Bar-plots depict number BGCs predicted in the genomes. **Panel-B**, the plot depicts the probability, P(I|n), of antagonistic activity or sensitivity by number of predicted BGCs. Increase in BGCs count is flowed by a decrease in the probability of high sensitivity and increase probability of antagonistic activity.

The bacterial root-microbiota members inhibit multi-drug resistant bacteria. Genomes and metabolomics analyses showed that root-associated bacteria are potentially eminent microorganisms to prospect for novel specialized metabolites. On the other hand, the ABBA screen has successfully showed that root-derived and abundant soil bacteria engage in antagonistic interactions. Moreover, the screen revealed also that most root-derived bacteria are high competitive members. Soil bacteria are well-known to produce clinically relevant antibiotics and may represent important reservoirs of novel antimicrobials (Tyc *et al.,* 2017). In order to extend the study of antagonistic interactions to human-related bacterial pathogens, we tested the potential of both root-derived and abundant soil bacteria to inhibit multi-drug resistant bacterial pathogens. All the 198 soil-derived bacteria have been tested three times independent for their antagonistic activity against several clinical bacterial isolates and against two reference strains, *Escherichia coli* and *Bacillus subtilis*, used during antibiotic susceptibility screens. The clinical isolates used in this study include six Gram-negative multi-drugs resistant bacteria, *Pseudomonas aeruginosa*, extended-spectrum β-lactamase (ESPL) and carbapenemase-producing (KPC) *Klebsiella pneumonia, Stenotrophomonas maltophilia, Acinetobacter baumanii* and *Heamophilus influenzae,* and six other Gram-positive multi-drugs resistant bacteria, *Staphylococcus aureus,* methicillin-resistant *Staphylococcus aureus, Staphylococcus epidermidis, Streptococcus pneumoniae, Streptococcus pyogens* and *Enterococcus feacium*. Among 198 root-derived and abundant soil bacteria, only 14 root-derived bacteria show antagonistic activity against several Gram-positive and Gram-negative bacteria (**Figure 18**). Among the 14 isolates that showed antagonism against multi-drug resistant bacteria, two isolates belong to *Actinobacteria*, one is a *Firmicutes* bacterium and eleven are *Proteobacteria* strains. These data indicate that root-associated bacteria have wider antagonistic activity and could be eminent candidates to screen for novel antibiotics.

The two *Actinobacteria* isolates that show antibiosis activity against multi-drugs resistant bacteria are *Streptomyces* n°1310 and *Nocardioidaceae* n°682 (**Figure 18**). Although both *Actinobacteria* isolates are Gram-positive bacteria, only the isolate n°682 could inhibit all other Gram-positive clinical isolates (**Figure 18**). This observation suggests that *Nocardioidaceae* bacteria are potentially good candidates to screen for antimicrobials that target other Gram-positive, although these bacteria are Gram-positive. All the clinical isolates in this study that are Gram-positive belong to the phylum *Firmicutes*. It is interesting to report that the isolates n°920 is the only *Firmicutes* bacterium from the culture collection that show antagonistic activity against Gram-positive and -negative clinical isolates (**Figure 18**). Moreover, despite isolate n°920 is a *Firmicutes* bacterium, the isolate could successfully inhibit phylum-related clinical isolates (**Figure 18**). However, the isolate n°920 appears as an exception since it is the only *Firmicutes* isolates that could inhibit multi-drugs resistant bacteria. Therefore, it is rather unlikely that root-derived *Firmicutes* bacteria could show antagonistic activity against multi-drugs resistant bacteria.

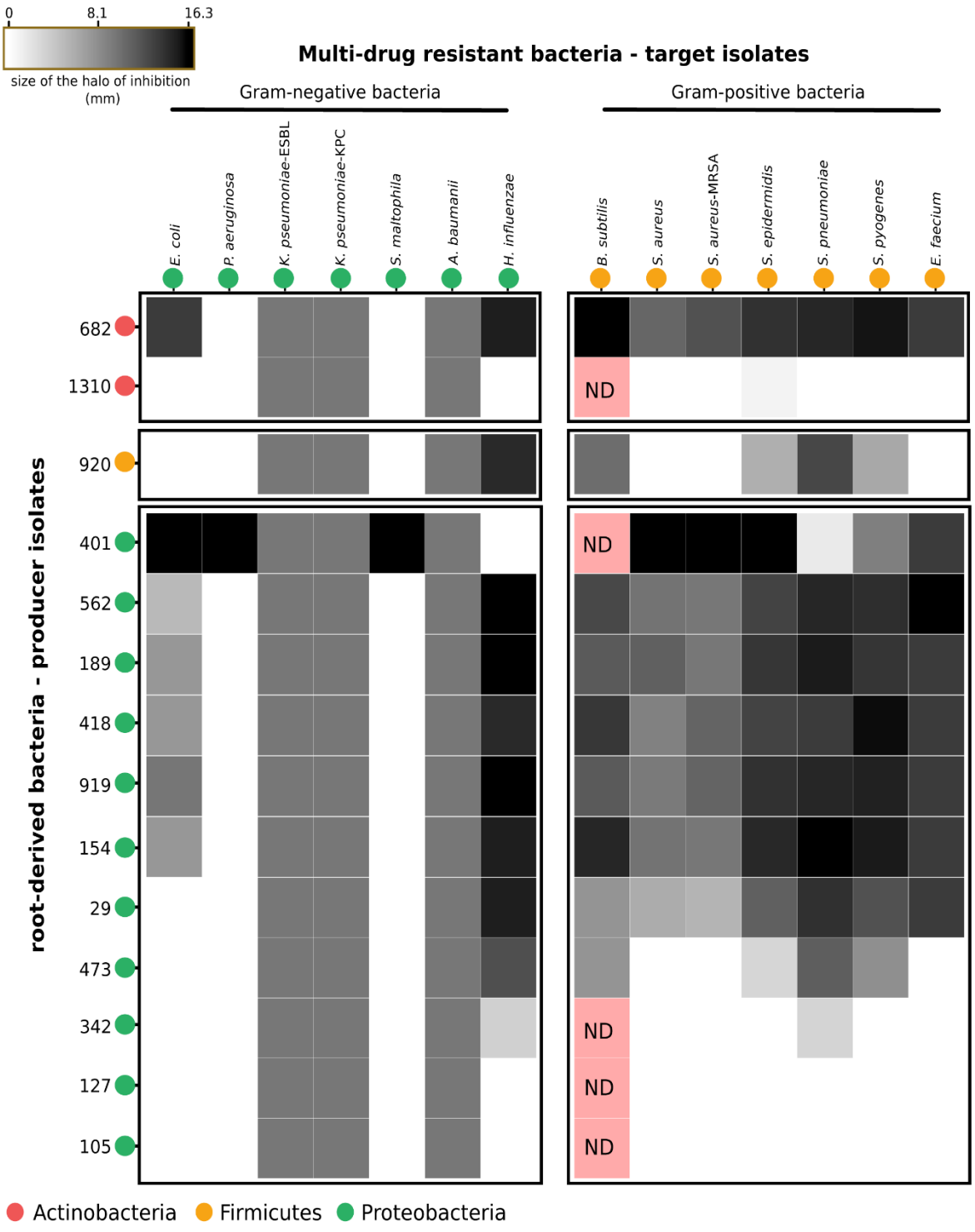


Figure 18| **Several *Arabidopsis* root microbiota members inhibit multi-drug resistant bacteria**

The heat map shows bacterial isolates that successfully inhibit different Gram-negative and Gram-positive clinical bacterial isolates. *E. coli* and *B. subtilis* are two reference strains used for antibiotic resistance test. Several root-derived bacteria, mainly Proteobacteria isolates, inhibit multi-drug resistant bacteria.

Most of the observed inhibitions against the clinical isolates are due to the antagonistic activity of *Proteobacteria* isolates (**Figure 18**). Remarkably, the *Pseudomonas* isolate n°401, which could antagonize 25% of the soil-derived bacteria, could also inhibit almost all tested clinical isolates (**Figure 18**). Even more interesting, the *Pseudomonas* strain n°401 is the only bacterium that could inhibit the

two Gram-negative pathogens *P. aeruginosa* and *S. maltophilia* (**Figure 18**). These data clearly indicate that high-competitive bacteria are more likely to inhibit niche-unrelated bacteria and are worth-prospecting to uncover novel antibiotics. Within the *Proteobacteria* isolates that inhibit multi-drugs resistant bacterial pathogens, two groups of isolates can be distinguished; a first one that inhibit all Gram-positive clinical isolates, and a second group that inhibit less or even none Gram-positive clinical isolates (**Figure 18**). The former group of isolates is composed by two *Pseudomonadaceae* n°401 and 562, two *Oxalobactereaceae* n°189 and 418, one *Burkholderiaceae*, *Sphingomonadaceae* and *Comamonadaceae* which are respectively n°919,154 and 29. The bacterial isolates in the latter group, n°473, 342, 127 and 105 belong to the following bacterial families, *Comamonadaceae*, *Caulobacteraceae*, *Rhizobiaceae* and *Hyphomicrobiaceae*, respectively. Interestingly, the former group is composed almost exclusively by Beta- and Gamma-*Proteobacteria*, except the isolates n°154 that is Alpha-*Proteobacteria* as most of the isolates in the latter group. Remarkably, Beta- and Gamma-*Proteobacteria* isolates seem to have more potent antibiosis activity than the Alpha-*Proteobacteria* against the Gram-positive pathogens (**Figure 18**). Taken all together, the screen for antagonistic activity against multi-drugs resistant clinical isolates indicates that the root-derived bacteria could secrete antimicrobials that inhibit several multi-drugs resistant bacteria. Moreover, the root-derived bacteria are eminent candidates to screen for novel antibiotics in order to combat the insurrection of multi-drugs resistant bacterial pathogens.

II.E. Discussion

Bacteria colonize a wide range of ecosystems where nutrients and space are limited. To compete against closely and distantly related species, bacteria have evolved a plethora of competitive mechanisms that help in providing resources and defining territoriality (Ghoul and Mitri, 2016). To a certain extent, bacteria with higher competitive potential may have fitness advantage over less competitive ones and could subsequently perpetuate their progeny. Insightful systematic analyses of genomes, metabolites and antagonistic activity are crucial to uncover the competitiveness potential of bacteria. Our multi-faceted approaches based on 1) comparative bacterial genomes analysis, 2) systematic identification of produced bacterial metabolites and 3) inter-bacterial antagonistic activity screen allowed us to identify the competitiveness potential of bacterial strains isolated from soil and *A. thaliana* roots. Mining the bacterial genomes revealed that root-derived and abundant soil bacteria harbor diverse BGCs, including clusters implicated in the biosynthesis of antimicrobial molecules like bacteriocins, microcins or several nonribosomal peptides or polyketides. Interestingly, the analysis of the bacterial metabolites has confirmed that some of the genome-predicted antimicrobials are produced by few bacteria when grown in a synthetic agar medium. These metabolites were either secreted by broadly or phylogenetically-restrained bacterial strains. More importantly, screen for mutual inhibitions has shown that bacteria engage in at-distance antagonistic interactions. Although antagonistic interactions are uncommon

among the bacterial isolates (2.5%), *Actinobacteria* in general and abundant soil bacteria in particular showed higher sensitivity to antagonistic activity. The combination of the three systematic analyses did not only help in exploring the competitive potential of each single isolate from the strain collection, but also provided a ground to define highly competitive and highly sensitive bacterial isolates. Defining groups of bacteria that have contrasting competitive potential is crucial for testing hypotheses that explore the role of these bacteria in altering the structure of microbial communities in different experimental systems.

As general rule, enzymatic pathways that produce diffusible antimicrobials are encoded in physically grouped set of genes known as biosynthetic gene clusters (BGCs) (Cimermancic *et al.,* 2014). The analysis of BGCs from the 198 draft genomes uncovered 1,404 clusters that cover 30 different classes. Interestingly, terpenes are by far the most commonly predicted class throughout the entire bacterial strains collection. Indeed, 80% of root-derived bacteria are predicted to harbor terpene, whereas only 50% of abundant soil bacteria are positive this class. These findings underline the importance of terpenes molecules for soil-derived bacteria and may furthermore suggest that these molecules are employed for long distant microbe-host or microbe-microbe interactions (Tyc *et al.*, 2017). While terpene gene cluster are widely predicted from the bacterial genomes, nonribosomal peptides (nrps) have been found to be the most abundant class within the isolates. This result is indicative that bacteria tend to possess more than one nrps cluster. Since nonribosomal peptides have broad biological functions (McIntosh *et al.,* 2010), it is more interesting to further investigate whether a same nrps gene cluster is conserved across several isolates. Interestingly, some BGC classes were predicted only within a small group of isolates. For instance the gene cluster for homoserine lactone or microcins are only pridected in *Proteobacteria* or *Firmicutes* isolates, respectively. Although the bacterial isolates are derived from soil, it is plausible that certain BGCs are adaptive traits that fit the species' specific niche requirements. Regardless the BGC class, it remains important to further investigate here whether a same biosynthetic gene cluster for each commonly predicted class is conserved across multiple distantly and/or closely related species. The additional genomic comparative analysis will shed light on which of horizontal or vertical genes transfers better explains how BGCs are spread and conserved across the strains.

Although this study is the first that explores BGCs encoded in the microbiota of *A. thaliana,* two similar studies have examined BGCs in 2,430 genomes from the human microbiota and 1,154 genomes of environmental bacteria (Cimermancic *et al.,* 2014, Donia *et al.,* 2014). In contrast to this study, the previous studies have independently indicated that saccharides gene clusters are by far the most abundant class in the human microbiota as well as in the environmental bacteria (Cimermancic *et al.,* 2014, Donia *et al.,* 2014). The discrepancy may be due to several reasons. First, draft genomes have been used in our study and not in the authors' studies where completed genomes have been used. Indeed, the assembly of BGCs is often fragmented and gaps are known to occur predominantly in large biosynthetic clusters, which prevent BGCs prediction by antiSMASH (Klassen and Currie, 2012). Second, in our work there was a taxonomic bias in the representative genomes across different

taxonomic levels and, more importantly, bias towards the cultivable fraction of the plant or soil microbiota. Including more *Bacteroidetes*, *Firmicutes* or metagenome samples is required to validate our conclusions and refine the difference observed between bacterial phyla. The third aspect is more related to the antiSMASH algorithm that fail to detect many classes of BGCs, among others oligosaccharide gene clusters (Cimermancic *et al.,* 2014). Including the algorithm "Cluster Finder" that detect BGCs independently of their class would increase the number and diversity of predicted BGCs from the culture collection as it has already been shown in the studies of Donia *et al.*,2014 and Cimermancic *et al.*,2014. In the light of these limitations, it is important to complete the genomes assembly, include more genome representatives and implement the algorithm "Cluster Finder" in order to better mine the genomes for BGCs within the plant microbiota. Matter of fact, the analysis of draft genomes from the soil-derived bacteria was also intended to reveal whether these isolates encode BGCs with predictive known antimicrobials in order to be further detected by analytic methods. However, it is important to highlight the recent meta-study of Tyc and colleagues that have explored BGCs content of 30 soil bacteria. The authors' study support the present one in the way that terpene, bacteriocin and nrps are abundantly predicted and oligosaccharides are almost undetectable (Tyc *et al.,* 2017). It is then plausible that bacteria from different ecosystems show different BGCs patterns and bacteria from comparable ecosystem show similar patterns in BGCs.

The antiSMASH analysis revealed that 8% of the overall predicted BGCs were functionally unknown and 92% of them covered 30 different classes. Several known antimicrobial molecules with broad-spectrum activity such as brabantamide (Schmidt *et al.,* 2014), phenazine (Borrero *et al.,* 2014), laspartomycin (Borders *et al.,* 2006) and rifamycin (Yu *et al.,* 1999), as well as narrow-spectrum antimicrobials like bacteriocines and microcins (Riley and Gordon, 1999), were predicted from the genomes. By analyzing the metabolites of all the bacteria grown separately in 25% tryptic soy agar, four different bioactive molecules have been found to be produced by a dozen of isolates. The staggering chemical diversity predicted by the genome mining approach could therefore not be recapitulated under laboratory conditions by LC-MS. The overlap between the prediction of BGCs *in silico* and the detection of the produced metabolites is a well-known phenomenon. Indeed, bacteria are known to harbor more BGCs than characterized molecules (Crüssmann *et al.,* 2016) and for several reasons a large majority of biosynthetic gene clusters are not expressed under laboratory conditions (Seyedsayamdost, 2014). In order to increase chances of detecting more diverse metabolites, including novel antimicrobials, it is important to vary the conditions where bacteria are grown, since different conditions might reflect different habitats. This assumption is supported by the study of Crüsemann *et al.,* 2016 where the authors could demonstrate that varying growth medium and solvent enhance the diversity of detectable metabolites. It is important to note that the metabolites identified through our LC-MS approach are constitutively produced and not induced upon challenge with competitor cells. Since antimicrobials are known to be employed by bacteria during inter-microbial warfare and that there production is tightly controlled in time and in space (Tyc *et al.,* 2014, Traxler *et al.,* 2013, Abrudan *et al.,* 2015), it is then

plausible that bacteria secrete more diverse set of antimicrobials upon challenge with a competitor. To test this hypothesis, we have initiated the description of secreted metabolites upon bacteria-bacteria interactions. Preliminary data indicate that bacteria diversify their metabolites upon competition sensing (*data not shown*). This preliminarily analysis indicates that identifying metabolites produced by a bacterium in co-culture with a competitor could lead to uncover new metabolites and antimicrobials. Although this study showed that only a few antibiotics are produced by bacteria, both genomes and metabolites analyses corroborate the assumption that *A. thaliana*- associated and/or soil-derived bacteria have the genetic and chemical potential to engage in antagonistic interactions at the soil-root interface.

Bacteria are known to secrete antimicrobial molecules and these have been anthropogenically exploited since decades to fight animal or plant microbial pathogens. Although an increasing number of studies have been exploring the antagonistic potential of bacteria to inhibit a multitude of pathogens, the extent to which antagonistic interactions occur between host-associated bacteria is still poorly studied. The screen for mutual inhibitions within and between root-derived and abundant soil bacteria revealed that antagonistic interactions are rather uncommon. Indeed, only 2.5% of overall tested interactions showed a halo of inhibition. The screen for mutual inhibitions between marine bacteria belonging to the family *Vibrionaceae* (Cordero *et al.*,2012) or for the inhibition of a Gram-positive and -negative bacteria by a large set of soil bacterial isolates (Tyc *et al.*,2014) corroborate this finding. In both former and latter studies, antagonistic interactions have been reported rather uncommon since only less than 6% of overall interactions are show a halo of inhibition. These data clearly indicate that bacteria do not frequently engage in the secretion of diffusible antimicrobials. Nonetheless, it is important to highlight that more than 65% of the isolates in the current study showed antagonistic activity against at least an isolate, suggesting that the large majority of root and soil-associated microbiota members employ antimicrobials for at-distance growth inhibition of specific competitors. Similarly, the study of antagonistic interactions between bacteria isolated from the medicinal plant *Echinaceae purpurea* (Maida *et al.,* 2015) or from the scleractinian corals *Montastrea annularis* (Rypien *et al.,* 2009) have also reported that a large proportion of bacteria show antagonistic activity against at least one bacterium. As a whole, these data constitute the evidence that a large proportion of environmental or host-associated bacteria are able to engage in at-distance competitive interactions and suggest that resistance to antibiosis is widespread. On the other hand, our study indicate that less than 40% of the isolates do not show at least once a sensitivity to antagonistic activity. Importantly, the results obtained by Maida *et al.,* 2015, Cordero *et al.,* 2012 and Rypien *et al.,* 2009 are in agreement with the finding that bacterial isolates that are resistant to all do not exceed 40% of the tested bacterial populations. Therefore, the statement that antagonistic interactions are uncommon among bacteria in mutual inhibition screens could be due to the fact that bacteria are often resistant to produced antimicrobials. Furthermore, resistance to antimicrobials could also be mediated by other mechanisms that involve competitive sensing (Cornforth and Foster, 2013). Indeed, bacteria can perform a chemical "camouflage" to prevent a competitor cell induce antibiotic productions, or bacteria can directly interfering in the production of antimicrobials (Abrudan *et al.*,

2015). In light of these assumptions, it is then important to consider screening for mutual inhibitions under antibiosis conducive conditions toward the attempt to disentangle between pure resistance to antibiosis and modulation of antagonistic activity.

The screen for mutual inhibitions among soil-derived bacteria has shown that several *Actinobacteria* isolates are highly sensitive to secreted antimicrobials. Moreover, *Actinobacteria* isolates showed also high intra-phylum (4%) and -family (11-22%) antagonistic activity. Although it is known that Gram-positive bacteria, including *Actinobacteria*, are more susceptible to antibiosis (Tyc *et al.,* 2014), it remains to be further investigated whether *Actinobacteria* from different habitat follow the same observed trend in sensitivity.

The bacterial culture collection used in this study included two groups of isolates that are ecologically delineated, root-associated bacteria and abundant soil bacteria. Interestingly, abundant soil bacteria have been reported here to be out-competed by the root-derived bacteria. This finding was partially expected since 68% of abundant soil bacteria are *Actinobacteria* isolates and *Actinobacteria* showed high sensitivity to antibiosis. Also, the root-derived bacteria are dominated by *Proteobacteria* isolates that are known to be resistant to antibiosis (Tyc *et al.,* 2014). Although a major limitation in the representative isolates among abundant soil bacteria, these data join the generally admitted assumption that fierce competitive interactions are to be expected more between than within bacterial populations (Pérez-Gutiérrez et al., 2013, Cordero *et al.,* 2012). Indeed, the results obtained by naturally occurring *Vibrionaceae* or between bacteria that have been isolated from different plant compartments strengthen the assumption that ecological competitions reflect to a certain extent the cohesion of a bacterial group. However, these data should be interpreted with caution since the scale of which bacteria interact is far smaller than the scale that is used to define these groups.

The analysis of the genomes, metabolites and the screen for mutual inhibitions allowed us to define distinct sets of *A. thaliana* and soil microbiota bacterial members with either highly competitive or sensitive phenotypes. Defining bacterial isolates with such contrasting competitive potential is primordial to generate testable hypotheses that explore the role of competition in altering the diversity and the structure of microbial communities in liquid microcosms and/or *in planta* by using gnotobiotic experimental systems (see **chapter III**).

II.F. References

- Abbamondi, G. R., De Rosa, S., Iodice, C., and Tommonaro, G. (2014). Cyclic dipeptides produced by marine sponge-associated bacteria as quorum sensing signals. *Nat Prod Commun* 9, 229–232.
- Abrudan, M. I., Smakman, F., Grimbergen, A. J., Westhoff, S., Miller, E. L., van Wezel, G. P., et al. (2015). Socially mediated induction and suppression of antibiosis during bacterial coexistence. *Proc Natl Acad Sci U S A* 112, 11054–11059. doi:10.1073/pnas.1504076112.
- Bai, Y., Müller, D. B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M., et al. (2015). Functional overlap of the Arabidopsis leaf and root microbiota. *Nature* 528, 364–369. doi:10.1038/nature16192.
- Borders, D. B., Leese, R. A., Jarolmen, H., Francis, N. D., Fantini, A. A., Falla, T., et al. (2007). Laspartomycin, an acidic lipopeptide antibiotic with a unique peptide core. *J. Nat. Prod.* 70, 443–446. doi:10.1021/np068056f.
- Borrero, N. V., Bai, F., Perez, C., Duong, B. Q., Rocca, J. R., Jin, S., et al. (2014). Phenazine antibiotic inspired discovery of potent bromophenazine antibacterial agents against Staphylococcus aureus and Staphylococcus epidermidis. *Org. Biomol. Chem.* 12, 881–886. doi:10.1039/c3ob42416b.
- Bulgarelli, D., Rott, M., Schlaeppi, K., Ver Loren van Themaat, E., Ahmadinejad, N., Assenza, F., et al. (2012). Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* 488, 91–95. doi:10.1038/nature11336.
- Cimermancic, P., Medema, M. H., Claesen, J., Kurita, K., Wieland Brown, L. C., Mavrommatis, K., et al. (2014). Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* 158, 412–421. doi:10.1016/j.cell.2014.06.034.
- Cordero, O. X., Wildschutte, H., Kirkup, B., Proehl, S., Ngo, L., Hussain, F., et al. (2012). Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. *Science* 337, 1228–1231. doi:10.1126/science.1219385.
- Cornforth, D. M., and Foster, K. R. (2015). Antibiotics and the art of bacterial war. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10827–10828. doi:10.1073/pnas.1513608112.
- Crüsemann, M., O'Neill, E. C., Larson, C. B., Melnik, A. V., Floros, D. J., da Silva, R. R., et al. (2017). Prioritizing Natural Product Diversity in a Collection of 146 Bacterial Strains Based on Growth and Extraction Protocols. *J. Nat. Prod.* 80, 588–597. doi:10.1021/acs.jnatprod.6b00722.
- Czárán, T. L., Hoekstra, R. F., and Pagie, L. (2002). Chemical warfare between microbes promotes biodiversity. *Proc Natl Acad Sci U S A* 99, 786–790. doi:10.1073/pnas.012399899.
- Davies, J., Spiegelman, G. B., and Yim, G. (2006). The world of subinhibitory antibiotic concentrations. *Curr. Opin. Microbiol.* 9, 445–453. doi:10.1016/j.mib.2006.08.006.

- Dombrowski, N., Schlaeppi, K., Agler, M. T., Hacquard, S., Kemen, E., Garrido-Oter, R., et al. (2017). Root microbiota dynamics of perennial Arabis alpina are dependent on soil residence time but independent of flowering time. *ISME J* 11, 43–55. doi:10.1038/ismej.2016.109.
- Donia, M. S., Cimermancic, P., Schulze, C. J., Wieland Brown, L. C., Martin, J., Mitreva, M., et al. (2014). A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. *Cell* 158, 1402–1414. doi:10.1016/j.cell.2014.08.032.
- Duquesne, S., Petit, V., Peduzzi, J., and Rebuffat, S. (2007). Structural and functional diversity of microcins, gene-encoded antibacterial peptides from enterobacteria. *J. Mol. Microbiol. Biotechnol.* 13, 200–209. doi:10.1159/000104748.
- Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N. K., Bhatnagar, S., et al. (2015). Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl. Acad. Sci. U.S.A.* 112, E911-920. doi:10.1073/pnas.1414592112.
- Federle, M. J., and Bassler, B. L. (2003). Interspecies communication in bacteria. *J. Clin. Invest.* 112, 1291–1299. doi:10.1172/JCI20195.
- Gerardin, Y., Springer, M., and Kishony, R. (2016). A competitive trade-off limits the selective advantage of increased antibiotic production. *Nat Microbiol* 1, 16175. doi:10.1038/nmicrobiol.2016.175.
- Ghoul, M., and Mitri, S. (2016). The Ecology and Evolution of Microbial Competition. *Trends Microbiol.* 24, 833–845. doi:10.1016/j.tim.2016.06.011.
- Ghoul, M., West, S. A., Johansen, H. K., Molin, S., Harrison, O. B., Maiden, M. C. J., et al. (2015). Bacteriocin-mediated competition in cystic fibrosis lung infections. *Proc. Biol. Sci.* 282. doi:10.1098/rspb.2015.0972.
- Guthals, A., Watrous, J. D., Dorrestein, P. C., and Bandeira, N. (2012). The spectral networks paradigm in high throughput mass spectrometry. *Mol Biosyst* 8, 2535–2544. doi:10.1039/c2mb25085c.
- Hayes, C. S., Koskiniemi, S., Ruhe, Z. C., Poole, S. J., and Low, D. A. (2014). Mechanisms and biological roles of contact-dependent growth inhibition systems. *Cold Spring Harb Perspect Med* 4. doi:10.1101/cshperspect.a010025.
- Hibbing, M. E., Fuqua, C., Parsek, M. R., and Peterson, S. B. (2010). Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol* 8, 15–25. doi:10.1038/nrmicro2259.
- Krug, D., and Müller, R. (2014). Secondary metabolomics: the impact of mass spectrometry-based approaches on the discovery and characterization of microbial natural products. *Nat Prod Rep* 31, 768–783. doi:10.1039/c3np70127a.
- Kusche, B. R., Phillips, J. B., and Priestley, N. D. (2009). Nonactin biosynthesis: setting limits on what can be achieved with precursor-directed biosynthesis. *Bioorg. Med. Chem. Lett.* 19, 1233–1235. doi:10.1016/j.bmcl.2008.12.096.

- Li, H., Lee, B.-C., Kim, T.-S., Bae, K.-S., Hong, J.-K., Choi, S.-H., et al. (2008). Bioactive Cyclic Dipeptides from a Marine Sponge-Associated Bacterium, Psychrobacter sp. *Biomolecules & Therapeutics* 16, 356–363. doi:10.4062/biomolther.2008.16.4.356.
- Maansson, M., Vynne, N. G., Klitgaard, A., Nybo, J. L., Melchiorsen, J., Nguyen, D. D., et al. (2016). An Integrated Metabolomic and Genomic Mining Workflow To Uncover the Biosynthetic Potential of Bacteria. *mSystems* 1. doi:10.1128/mSystems.00028-15.
- Maida, I., Chiellini, C., Mengoni, A., Bosi, E., Firenzuoli, F., Fondi, M., et al. (2016). Antagonistic interactions between endophytic cultivable bacterial communities isolated from the medicinal plant Echinacea purpurea. *Environ. Microbiol.* 18, 2357–2365. doi:10.1111/1462-2920.12911.
- McIntosh, J. A., Robertson, C. R., Agarwal, V., Nair, S. K., Bulaj, G. W., and Schmidt, E. W. (2010). Circular logic: nonribosomal peptide-like macrocyclization with a ribosomal peptide catalyst. *J. Am. Chem. Soc.* 132, 15499–15501. doi:10.1021/ja1067806.
- Medema, M. H., Cimermancic, P., Sali, A., Takano, E., and Fischbach, M. A. (2014). A systematic computational analysis of biosynthetic gene cluster evolution: lessons for engineering biosynthesis. *PLoS Comput. Biol.* 10, e1004016. doi:10.1371/journal.pcbi.1004016.
- Peiffer, J. A., Spor, A., Koren, O., Jin, Z., Tringe, S. G., Dangl, J. L., et al. (2013). Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6548–6553. doi:10.1073/pnas.1302837110.
- Raaijmakers, J. M., and Mazzola, M. (2012). Diversity and natural functions of antibiotics produced by beneficial and plant pathogenic bacteria. *Annu Rev Phytopathol* 50, 403–424. doi:10.1146/annurev-phyto-081211-172908.
- Pérez-Gutiérrez, R.-A., López-Ramírez, V., Islas, Á., Alcaraz, L. D., Hernández-González, I., Olivera, B. C. L., et al. (2013). Antagonism influences assembly of a Bacillus guild in a local community and is depicted as a food-chain network. *ISME J* 7, 487–497. doi:10.1038/ismej.2012.119.
- Riley, M. A., and Gordon, D. M. (1999). The ecological role of bacteriocins in bacterial competition. *Trends Microbiol.* 7, 129–133.
- Rutledge, P. J., and Challis, G. L. (2015). Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat. Rev. Microbiol.* 13, 509–523. doi:10.1038/nrmicro3496.
- Rypien, K. L., Ward, J. R., and Azam, F. (2010). Antagonistic interactions among coral-associated bacteria. *Environ. Microbiol.* 12, 28–39. doi:10.1111/j.1462-2920.2009.02027.x.
- Sassone-Corsi, M., Nuccio, S.-P., Liu, H., Hernandez, D., Vu, C. T., Takahashi, A. A., et al. (2016). Microcins mediate competition among Enterobacteriaceae in the inflamed gut. *Nature* 540, 280–283. doi:10.1038/nature20557.

- Schmidt, Y., van der Voort, M., Crüsemann, M., Piel, J., Josten, M., Sahl, H.-G., et al. (2014). Biosynthetic origin of the antibiotic cyclocarbamate brabantamide A (SB-253514) in plant-associated Pseudomonas. *Chembiochem* 15, 259–266. doi:10.1002/cbic.201300527.
- Seyedsayamdost, M. R. (2014). High-throughput platform for the discovery of elicitors of silent bacterial gene clusters. *Proc. Natl. Acad. Sci. U.S.A.* 111, 7266–7271. doi:10.1073/pnas.1400019111.
- Stubbendieck, R. M., and Straight, P. D. (2016). Multifaceted Interfaces of Bacterial Competition. *J. Bacteriol.* 198, 2145–2155. doi:10.1128/JB.00275-16.
- Stubbendieck, R. M., Vargas-Bautista, C., and Straight, P. D. (2016). Bacterial Communities: Interactions to Scale. *Front Microbiol* 7, 1234. doi:10.3389/fmicb.2016.01234.
- Subramanian, S., and Smith, D. L. (2015). Bacteriocins from the rhizosphere microbiome –from an agriculture perspective. *Front Plant Sci* 6, 909. doi:10.3389/fpls.2015.00909.
- Traxler, M. F., Watrous, J. D., Alexandrov, T., Dorrestein, P. C., and Kolter, R. (2013). Interspecies interactions stimulate diversification of the Streptomyces coelicolor secreted metabolome. *MBio* 4. doi:10.1128/mBio.00459-13.
- Tyc, O., Song, C., Dickschat, J. S., Vos, M., and Garbeva, P. (2017). The Ecological Role of Volatile and Soluble Secondary Metabolites Produced by Soil Bacteria. *Trends Microbiol.* 25, 280–292. doi:10.1016/j.tim.2016.12.002.
- Tyc, O., van den Berg, M., Gerards, S., van Veen, J. A., Raaijmakers, J. M., de Boer, W., et al. (2014). Impact of interspecific interactions on antimicrobial activity among soil bacteria. *Front Microbiol* 5, 567. doi:10.3389/fmicb.2014.00567.
- von Wintersdorff, C. J. H., Penders, J., van Niekerk, J. M., Mills, N. D., Majumder, S., van Alphen, L. B., et al. (2016). Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. *Front Microbiol* 7. doi:10.3389/fmicb.2016.00173.
- Wang, H., Fewer, D. P., Holm, L., Rouhiainen, L., and Sivonen, K. (2014). Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. *Proc. Natl. Acad. Sci. U.S.A.* 111, 9259–9264. doi:10.1073/pnas.1401734111.
- Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Bruccoleri, R., et al. (2015). antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* 43, W237-243. doi:10.1093/nar/gkv437.
- Yu, T.-W., Shen, Y., Doi-Katayama, Y., Tang, L., Park, C., Moore, B. S., et al. (1999). Direct evidence that the rifamycin polyketide synthase assembles polyketide chains processively. *Proc Natl Acad Sci U S A* 96, 9051–9056.

- Ziemert, N., Podell, S., Penn, K., Badger, J. H., Allen, E., and Jensen, P. R. (2012). The Natural Product Domain Seeker NaPDoS: A Phylogeny Based Bioinformatic Tool to Classify Secondary Metabolite Gene Diversity. *PLoS One* 7. doi:10.1371/journal.pone.0034064.

Perturbation by *in vivo* depletion of community members' in order to study the role of bacteria-bacteria interactions in the establishment of microbial communities in liquid microcosms and *in planta*

III.A. Introduction

Healthy and asymptomatic plants are colonized by an astonishing diversity of microorganisms that are collectively known as the plant microbiota (Müller *et al.,* 2016). Bacteria, major component of the plant microbiota, colonize both above- and below-ground plant's tissue (Bulgarelli *et al.,* 2013, Vorholt, 2012) and provide fitness advantages to the host *via* various mechanisms (Friesen *et al.,* 2011). For instance, the bacterial microbiota can help the plant to cope abiotic stresses (Yang *et al.,* 2008) or can provide a protective shield against pathogenic microbial invaders through colonization of available space and/or secretion of antimicrobials (Whipps, 2001). The advent of high-throughput 16S rDNA amplicon sequencing has permitted the study of microbial communities associated with diverse plant ecotypes grown under various conditions (Dombrowski *et al.,* 2017, Wagner *et al.,* 2016, Edwards *et al.,* 2015, Schlaeppi *et al.,* 2013). While our knowledge on the composition and function of the plant microbiota is exponentially expending, our understanding of the fundamental principles and assembly rules that govern plant microbiota establishment are still elusive. Particularly, the role of microbe-microbe interactions for structuring and stabilizing microbial networks along the plant-root continuum is not well understood. The lack of in-depth studies that explore the role of bacteria-bacteria competitive or cooperative interactions in the establishment of host-associated microbiota is due to, out of many other factors, the complexity of microbial communities associated with the plants. In order to deconvolute microbiota complexity, it is useful to work with simplified synthetic microbial communities, yet representative of the plant microbiota, and to employ gnotobiotic systems that partly mimic natural habitat but under strictly controlled laboratory conditions. In the recent years, several studies have elegantly utilized synthetic bacterial communities that are representative of the plant microbiota to demonstrate the role of plant genotypes, immune system or plant tissue organs in altering the structure of bacterial communities (Bai *et al.,* 2015, Lebeis *et al.,* 2015, Bodenhausen *et al.,* 2014). More recently, the study of simplified seven-species community members of the maize root microbiota have highlighted the role of bacteria-bacteria interactions in the assembly of these bacteria on the host's roots under axenic growth condition (Niu *et al.,* 2017). Interestingly, while it is known that selection, drift, speciation and dispersal are four processes that govern the establishment of plant-associated bacterial communities (Herrera Paredes and Lebeis, 2016), the extent to which bacteria-bacteria competitive interactions alter the composition and the structure of the plant root microbiota is still not well studied.

In this chapter, I aim at studying the role of bacteria-bacteria interactions in the establishment of bacterial communities *in vitro* and *in planta*. To this end, I used the same bacterial culture collection described in chapter II and performed perturbation by community member depletion experiments. Based on systematic analyses of the bacterial competitive potential presented in the previous chapter, I test the

hypothesis that *in vivo* depletion of the 13 highly competitive bacterial strains alters more strongly the community diversity and structure than the *in vivo* depletion of the 13 highly sensitive bacteria (see chapter II, part C). To test the aforementioned hypothesis, I used two closed experimental systems: an *in vitro* liquid microcosms and an *in planta* gnotobiotic systems. The former is comprised of two liquid bacterial growth media (minimal and complex) under two growth states (standing and shaking). The second experimental system consists of a solid matrix-like soil that is either depleted of complex organic matter (designated by calcined clay) or amended with 3% complex organic matter in the form of peat (designated by calcined clay plus 3% peat). These two different matrices are used for microbiota reconstitution experiments of germ-free *Arabidopsis thaliana*. Interestingly, *in vivo* depletion of the 13 highly competitive strains led to a sharp decrease in species richness of microbial communities and strongly altered the community structure in liquid microcosms. In contrast, *in vivo* depletion of the 13 highly sensitive bacteria had a marginal effect on the community diversity and structure within comparable liquid microcosms. Although counter-intuitive, these data indicate the important role of competition in promoting community diversity and stability and are supported by theory-based studies (Coyte *et al.,* 2016, Vetsigian *et al.,* 2011, Czárán *et al.* 2001). Further analysis of *in planta* synthetic bacterial communities indicated that the *Arabidopsis thaliana* root-associated microbiota is resilient to perturbation by community member depletions. Jointly, these data indicate that highly competitive community members are important for the promotion of community diversity and stability in a niche-dependent manner.

III.B. *In vivo* depletion of highly competitive strains alters strongly species richness and community structure in liquid microcosms

Distinct bacterial communities assemble under different liquid microcosms. A large proportion of root-associated bacteria are derived from the surrounding soil (Bulgarelli *et al.,* 2012). Soil is known to be poor in nutrients compare to the vicinity of the roots where plants exude surplus carbon sources (Fierer *et al.,* 2007). Therefore, soil-derived root-associated bacteria transit from nutrient-poor to nutrient-rich habitat. Although the root vicinity is carbon-rich, higher bacterial diversity is observed in bulk soil than in the rhizosphere (area adjacent to the roots) or rhizoplane (root surface) (Bulgarelli *et al.,* 2012). Moreover, the structure of microbial communities in the bulk soil are distinct from those of the rhizosphere or the rhizoplane (Bulgarelli *et al.,* 2012, Lundberg *et al.,* 2012). Herein, I test whether contrasting bacterial growth conditions (liquid microcosms) affect the community structure. To this end, I incubated the 198 isolates studied in the preceding chapter in four growth conditions: minimal or complex medium and two states shaking or standing (**Figure 19, panel-A**). After 96h of incubation at 25°C, output communities were analyzed using amplicon sequencing of the regions V5-V7 of the 16S rRNA gene. Out of 198 inoculated strains, 130 isolates can be distinguished at 100% sequence similarity

according to V5-V7 regions of the 16S rRNA gene. Bacterial isolates that cannot be distinguished are de-replicated within a unique representative sequence (**Figure 20**).

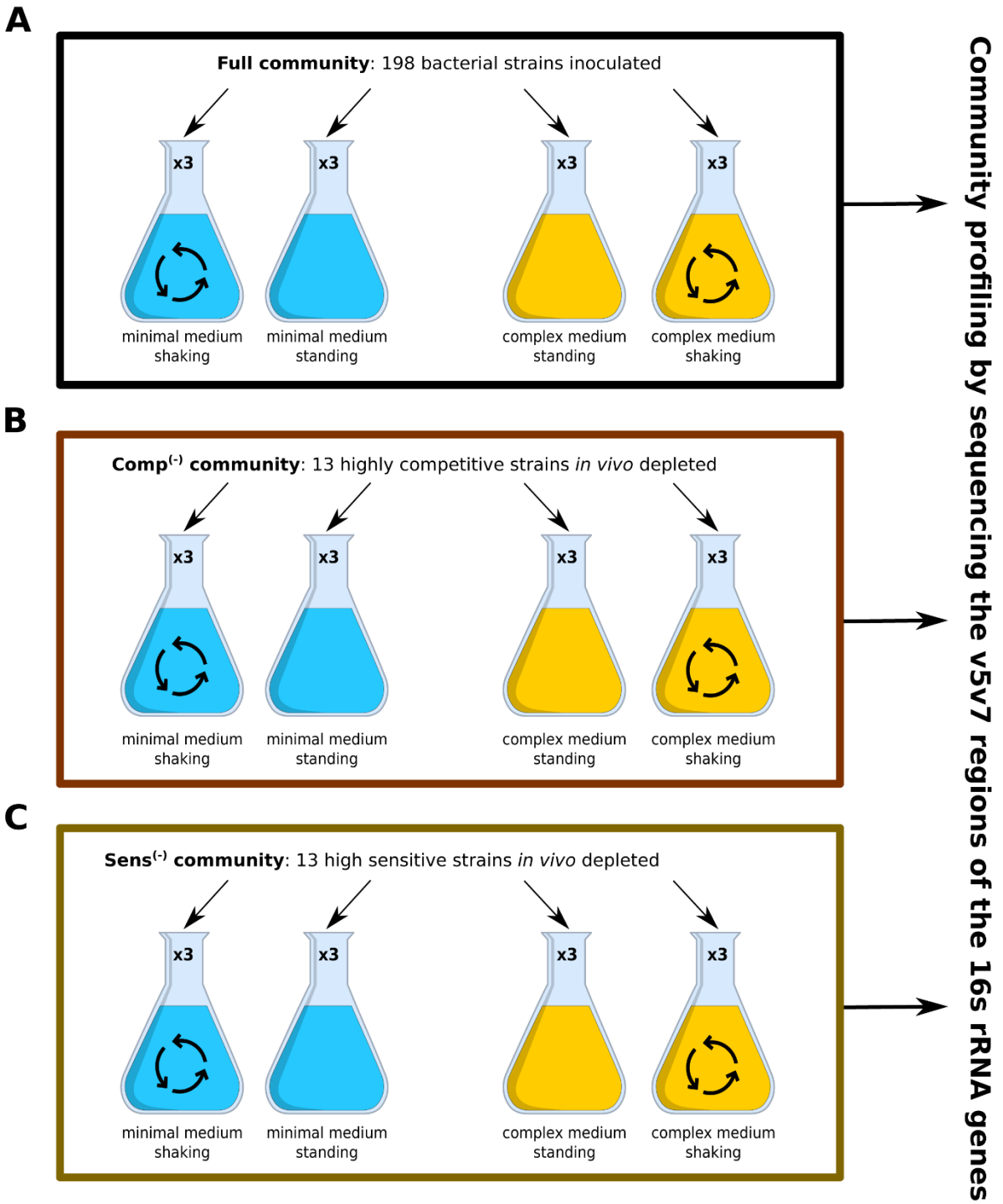


Figure 19| **Experimental setup of *in vitro* perturbation by community members' depletion experiment.**

Flowcharts depicting experimental design of perturbation by community members' depletion experiment. Tested conditions include two different media, minimal and complex, that are color coded, and two growth states, standing and shaking. Shaking state is indicated by three arrows forming a circle. Each condition include three biological replicates and each biological replicate have two technical replicates. **Panel-A**, 198 bacterial isolates are inoculated in four different microcosms and referred by full community. **Panel-B**, 13 highly competitive strains are *in vivo* depleted from the full community and referred by Comp$^{(-)}$ community. **Panel-C**, 13 highly sensitive strains are *in vivo* depleted from the full community and referred as Sens$^{(-)}$ community. Bacterial communities were incubated for 96h at 25°C. Input and output communities were profiled by sequencing v5v7 regions of the 16s rRNA genes.

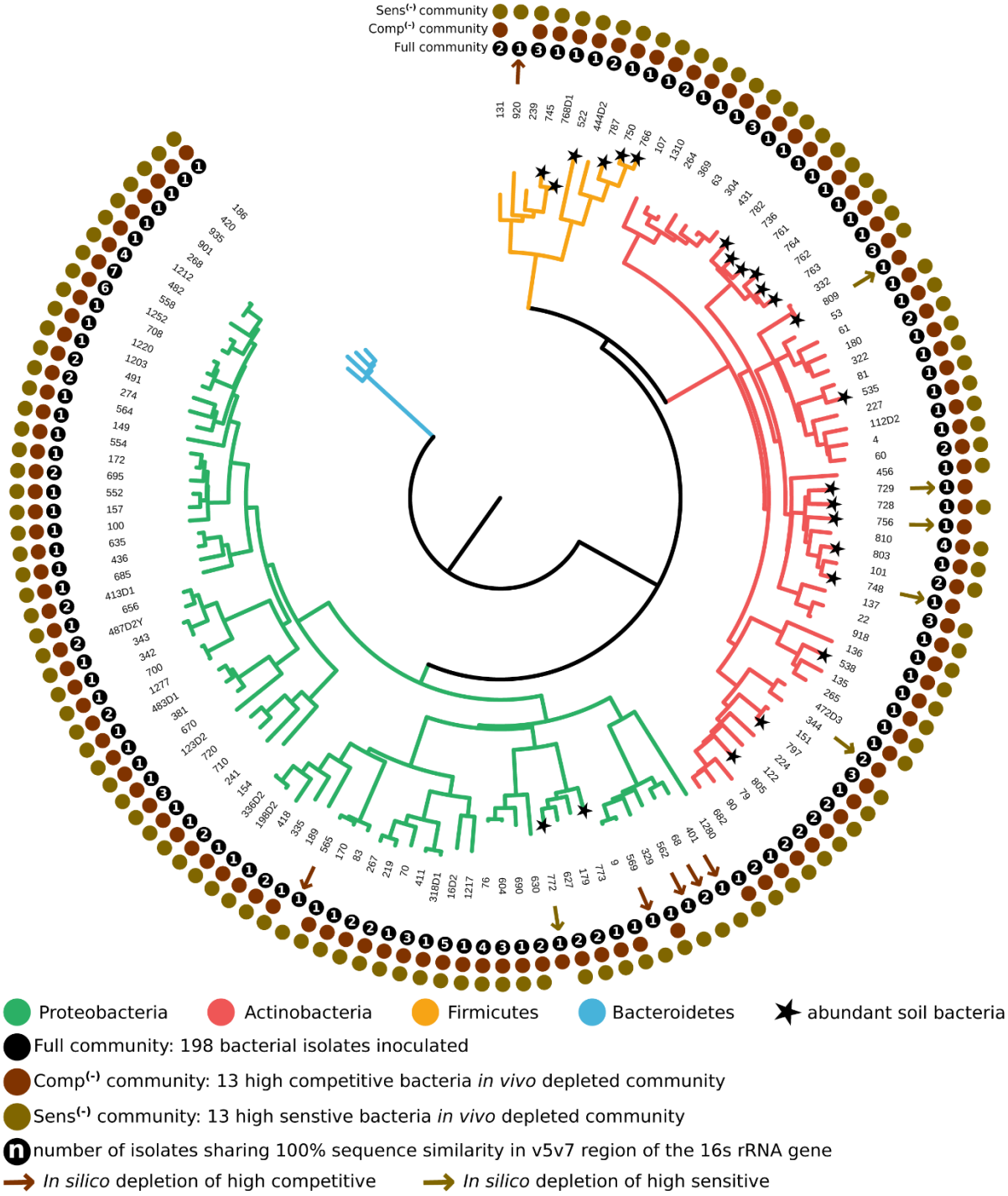


Figure 20| **Phylogenetically diverse bacteria are employed in perturbation by community members' depletion experiment.**

Circular phylogenetic tree shows 16s rRNA genes distances between bacterial strains used in the perturbation by community members' depletion experiment. The phylogenetic tree corresponds to neighbor joining tree built from multiple alignment of the v5v7 regions of the 16s rRNA genes. The color code in branches indicates phylum and stars indicate abundant soil bacteria. Circle at a leaf indicates inoculated isolate(s) and color in circle refers to one of the three synthetic communities; 1)- full community, Comp$^{(-)}$ community, or Sens$^{(-)}$ community. Values indicated inside black circles refer to the number of de-replicated sequences within each representative sequence. Arrow points to *in silico* depleted isolates and color in arrows indicate highly competitive or highly sensitive strains.

To examine whether the tested growth conditions alter the bacterial community composition, I computed two different alpha-diversity measures (Shannon index and Observed species). The former measures species richness and evenness, whereas the latter only takes into account the number of observed species. Both methods were applied on rarefied to even sequencing depth which corresponds to the smallest sample size of the species count data. The analysis of the alpha-diversity from both input (strains mixed at T_0) and output communities (formed in the liquid microcosms after 96 hours of incubation) indicates a significant and sharp drop in species richness in output communities (**Figure 21 panel A** and **B**). The decrease in species richness in output communities might be due to the extinction of species that could be a consequence of competitive inter-bacterial interactions. Interestingly, minimal medium in shaking state maintains higher alpha-diversity than complex medium in a similar growth state. This observation holds also true for standing microcosms (**Figure 21**). These data indicate that maximum diversity is to be expected under growth-limiting conditions rather than under nutrient-rich conditions. It is plausible that under growth-limiting conditions bacteria regulate the production of antimicrobials that have high energetic cost (Aguirre-von-Wobeser *et al.*, 2015). The decrease in species richness in complex medium is to some extent reminiscent of differences in species richness observed between rhizosphere (*i.e.* carbon-rich) and bulk soil (*i.e.* carbon-poor). Interestingly, the total number of observed species between complex medium in shaking and minimal medium in standing are comparable. However, former and letter conditions show a significant difference in Shannon index (**Figure 21**). Higher Shannon index in shaking microcosms indicates that more species are evenly distributed across the replicates. These data suggest that bacterial communities in shaking growth conditions are expected to be more homogeneous than communities in standing microcosms.

To test whether bacterial communities in shaking cultures are significantly more homogeneous than communities in standing cultures, I measured Bray Curtis (BC) distance to the centroid and within group distance (between samples belonging to a same grouping factor) for each output community. The grouping factor corresponds to the four growth conditions indicated above and encompasses all biological and technical replicates. The analysis of group homogeneity reveals a significant decrease in BC distances to centroid as well as within group distance for both tested mediums in shaking (**Figure 22**). These data indicate that microbial communities assembled in shaking microcosms are more homogeneous than those assembled in standing microcosms. Nutrients and oxygenation are presumably more uniformly distributed in agitated microcosms than in non-agitated. In contrast, a standing condition allows the formation of several micro-habitats within the microcosm. It is therefore more likely that observed high dispersal in the structure of bacterial communities grown in a standing condition is caused by a stochastic colonization of several micro-habitats within the microcosm. Interestingly, bacterial communities grown in minimal medium tend to be more homogeneous than communities grown in complex medium (**Figure 22**). These data indicate that bacterial assemblages in a complex medium are more subject to stochastic colonization than in a minimal medium. Observed stochasticity in the

colonization of complex medium could be linked, among other factors, to priority effect and inter-bacterial competitive interactions.

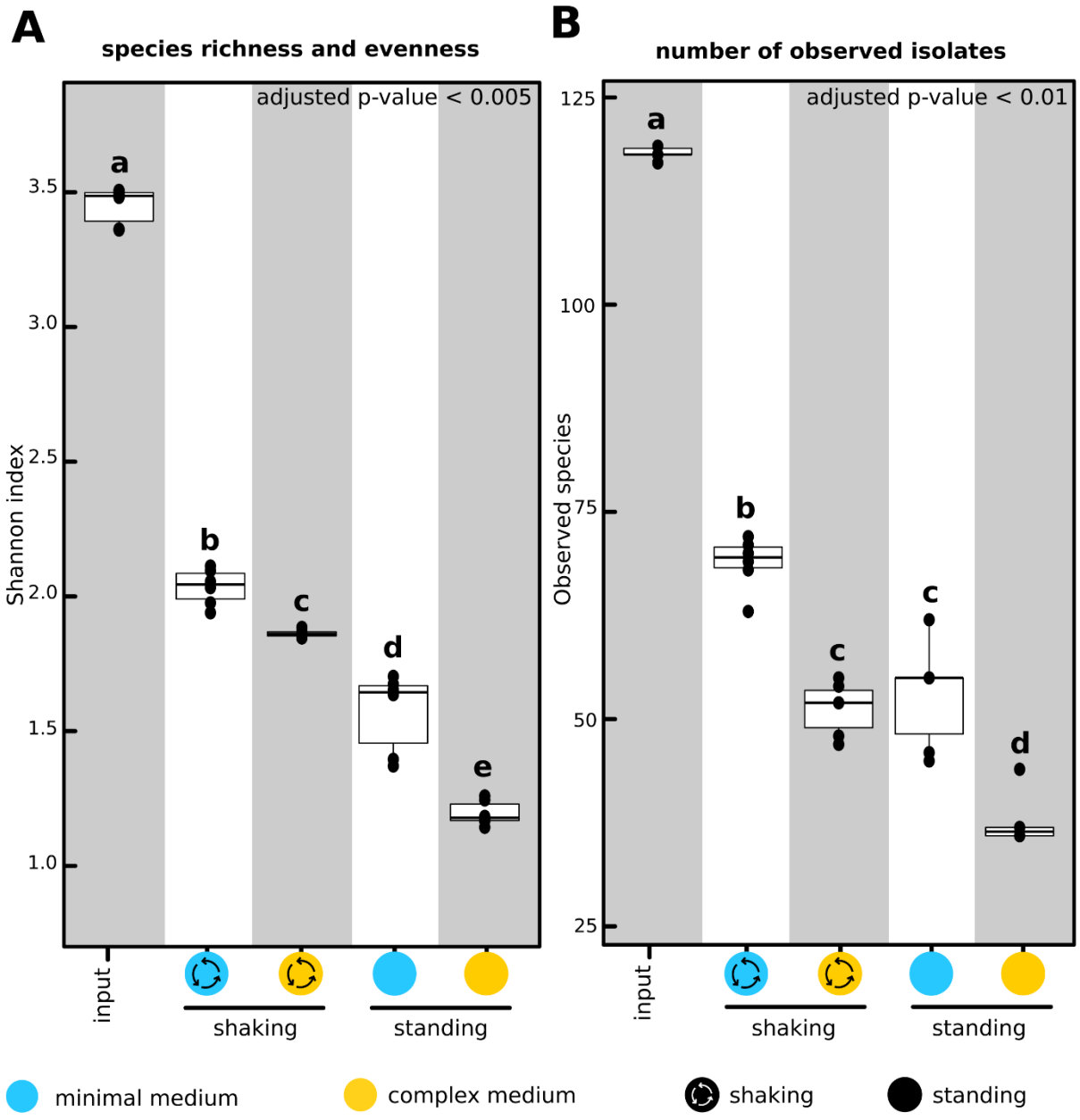


Figure 21| **Tested growth conditions alter species richness and evenness.**

The boxplots depict alpha-diversity of input and output communities. **Panel-A**, the Shannon index measure of the alpha-diversity indicates species richness and evenness across the four microcosms and input communities. **Panel-B**, the Observed species measure indicates number of isolates retrieved from input and output communities. All output community show significant decrease in species richness compared to the input. Minimal medium shows significantly higher Shanon index and Observed species than minimal medium under shaking or standing state. Observed species in minimal or complex medium under shaking condition are significantly higher than in standing condition, except shaking condition in complex medium that is not significantly different from standing condition in minimal medium.

To further investigate the structure of bacterial communities assembled in the four growth conditions, I performed a principal coordinate analysis (PCoA) on BC distances between all samples. The analysis is illustrated in **Figure 23, panel-A** (unconstrained PCoA) and in **Figure 23, panel-B** (constrained PCoA). Samples belonging to the same growth condition share same color and shape code. The unconstrained analysis reveals that growth medium, minimal or complex, explains most of the variance (up to 46%) and growth state, shaking or standing, explains up to 15% of the variance (**Figure 23, panel-A**). In order to test the significance of observed variances, I constrained the analysis by growth medium in the first axis and by growth state in the second axis. The growth medium explains 40.9% of the variance with a corrected *p*-value of 0.001. In the counterpart, the growth state explains only 12.4% of the variance with a corrected *p*-value of 0.001 (**Figure 23, panel-B**). Accordingly, medium composition explains significantly more of the variance than growth state. More importantly, these data show that the four growth conditions lead to the assembly of distinct bacterial communities. Although all the growth conditions share the same input communities, output communities under tested conditions are significantly different by their species composition and community structure.

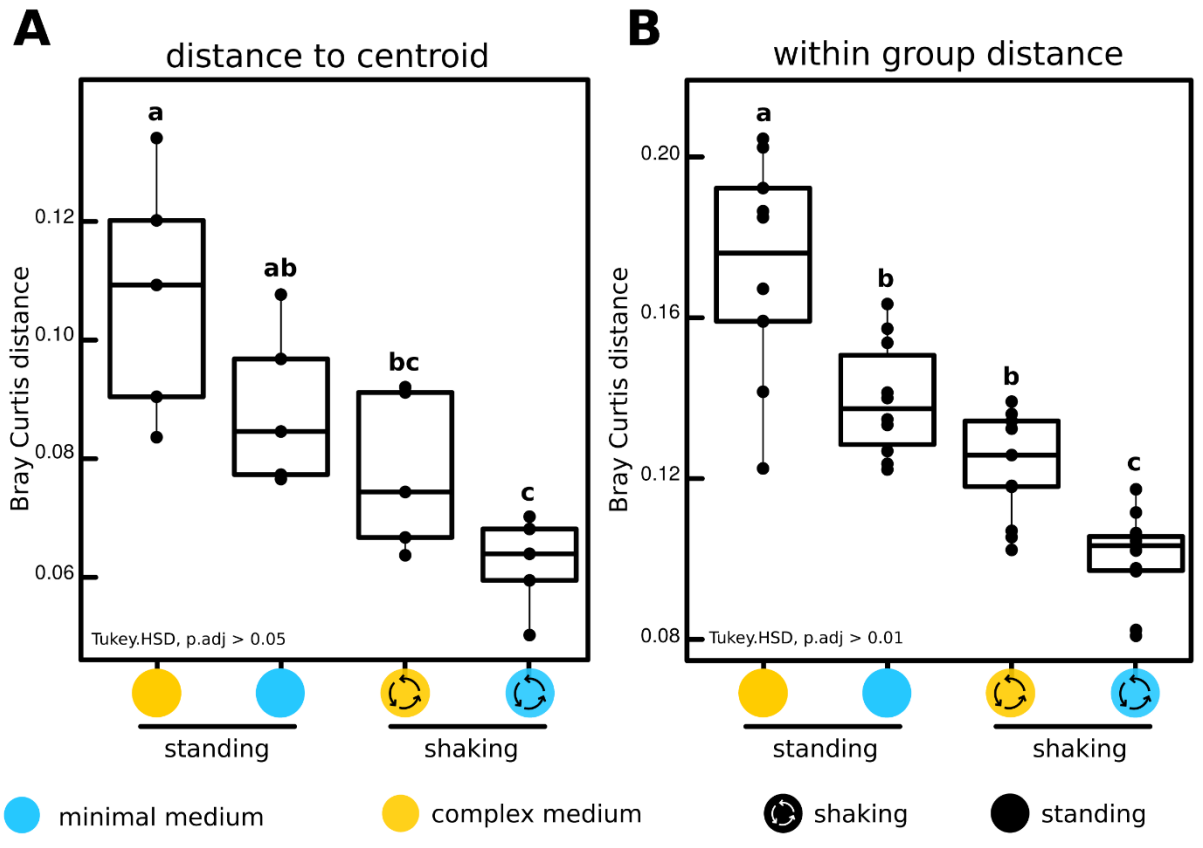


Figure 22| **Applied growth conditions alter the homogeneity of bacterial communities.**

Panel-A and **B** depict respectively Bray Curtis (BC) distance to centroid and within group BC distance. The interaction between growth medium and growth state correspond to one grouping factor and represent all replicates of one microcosm. Bacterial communities under shaking state in minimal or complex medium show significantly low distance to centroid and within group distance than communities under standing state in similar growth medium. Communities in minimal medium under shaking state are significantly more homogeneous than all other communities in tested microcosms.

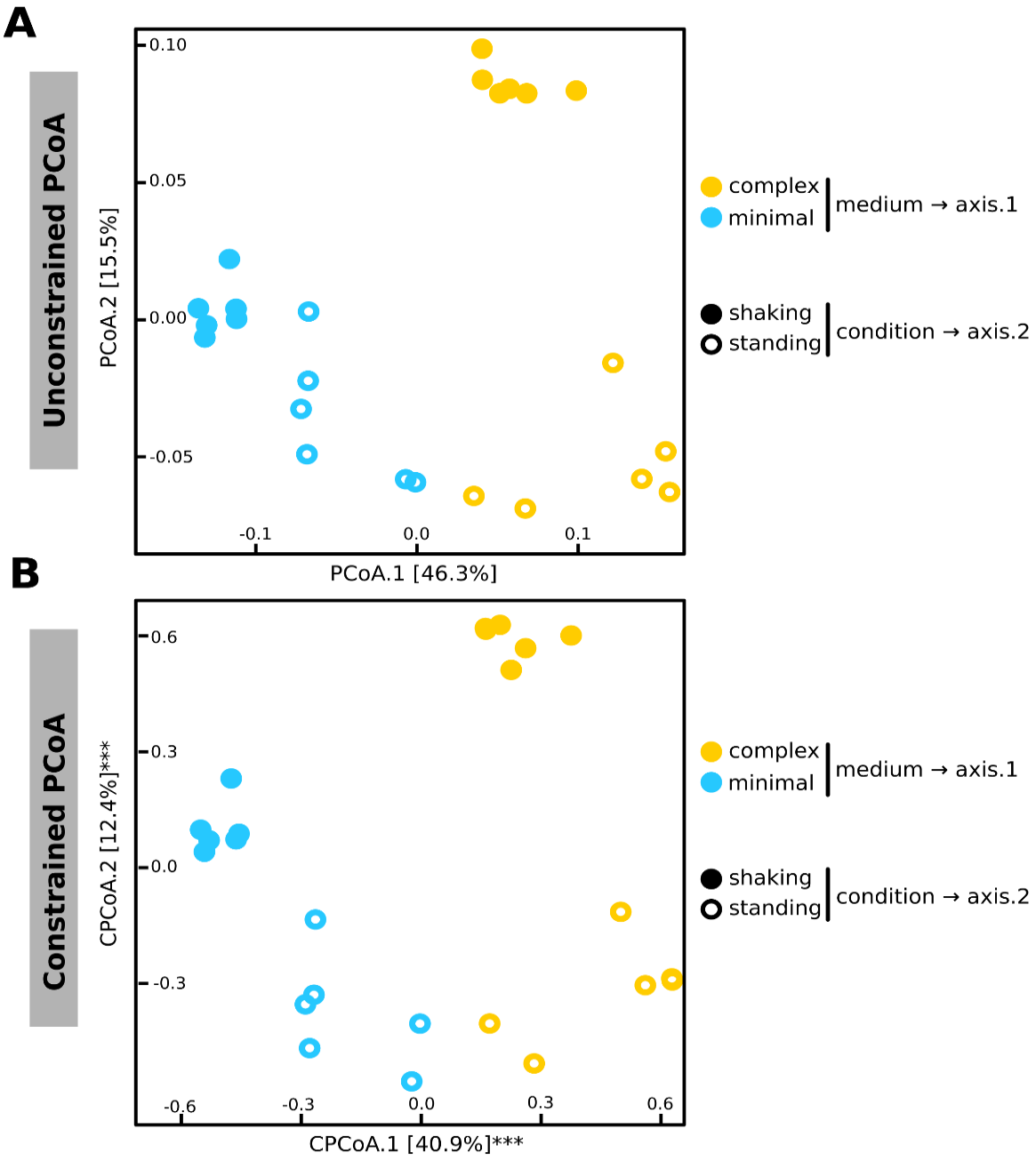


Figure 23| **Distinct bacterial communities assemble under different microcosms.**

Panel-A and **B** depict respectively unconstrained and constrained Principal Coordinates Analysis "PCoA" on Bray Curtis (BC) distances between bacterial communities under different growth conditions. Each shape represents a bacterial community (one sample). Color code indicates medium and filled or unfilled shapes indicate shaking or standing state, respectively. Distances between samples correspond to BC dissimilarity distance computed on normalized count data. **Panel-A**, first axis of the PCoA explains most of the observed variance in the data and separate the samples according to growth medium. Second axis in the plot separates the samples according to growth state. **Panel-B**, constrained analysis indicate that growth medium explains up to 40.9% of the variance with *p-value* of 0.001 and growth state explains 12.4% of the variance with *p-value* of 0.001.

The observed differences between tested microcosms point to the preferential enrichment of certain isolates under a specific condition. In order to reveal the bacterial isolates that contribute most to the observed shift in the community structure, I performed an enrichment test by comparing the $\log_2$ fold change in the relative abundance of the isolates between; minimal medium shaking *versus* complex medium shaking and between minimal medium standing *versus* complex medium standing. The results of the enrichment tests are depicted in **Figure 24**. Interestingly, the analysis of the $\log_2$ fold change in

the relative abundance shows that several bacterial isolates belonging to diverse *Proteobacteria* families, but mainly *Burkholderiales* and *Rhizobiales*, and to several *Actinobacteria* families, but mainly *Microbacteriaceae, Micrococcaceae*, *Intrasporangiaceae* and *Promicromonosporaceae,* are significantly enriched in minimal medium (upper part right side, **Figure 24**). In the contrary, only a few isolates are significantly enriched in complex medium, and they are predominantly *Pseudomonadaceae, Xanthomonadaceae, Bacillaceae* or *Streptomycetaceae* (lower part left side, **Figure 24**). These data indicate that fewer strains, and mainly from the γ-*Proteobacteria* class, are significantly enriched in complex medium. All together, these data clearly show that all tested growth conditions lead to the assembly of distinct bacterial communities. These communities are different in their species composition, community structure and preferentially enriched isolates. It is therefore important to examine the effect of *in vivo* depletion of community members on the community richness and structure in these distinct microcosms irrespectively of the shared medium or state.

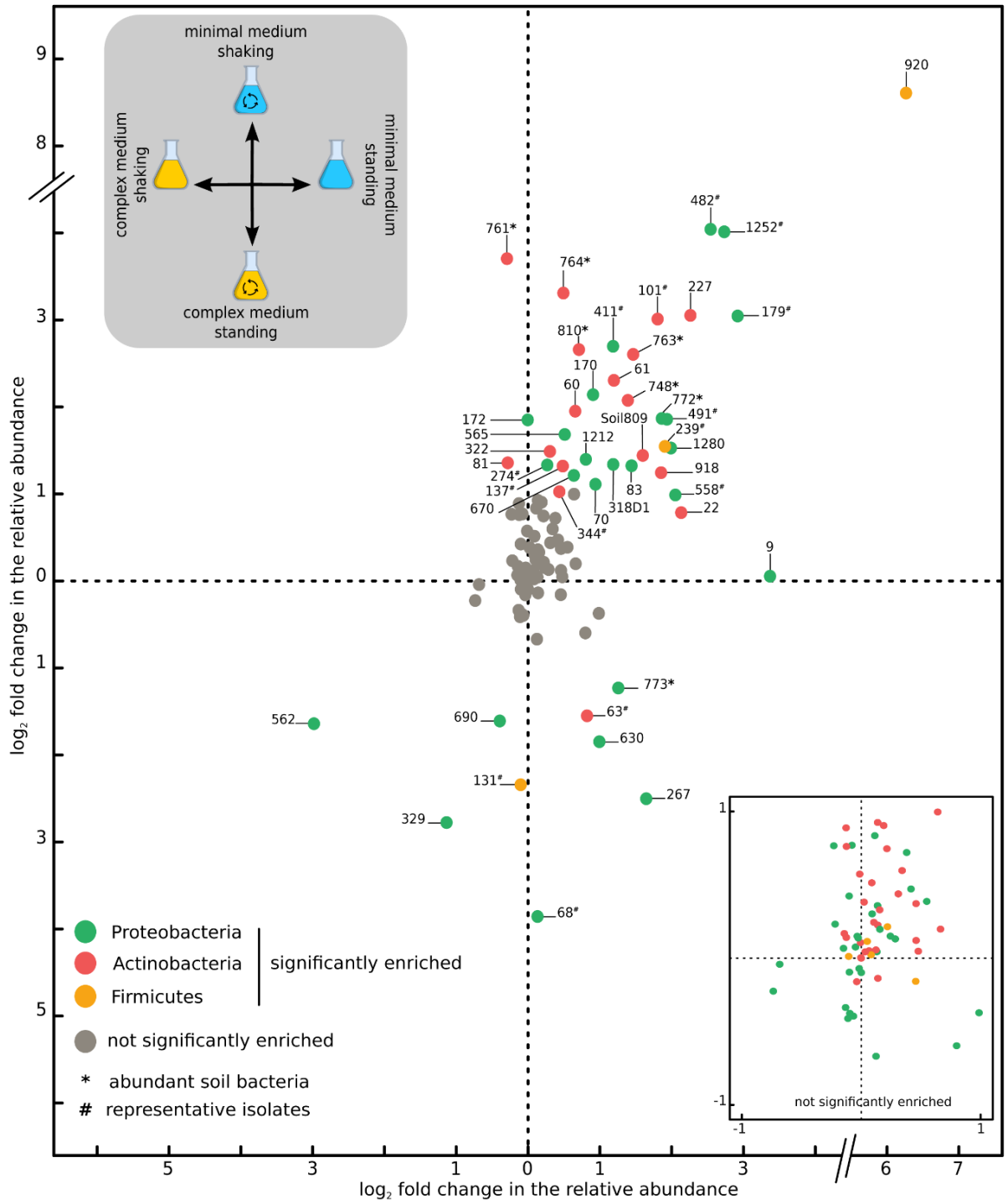


Figure 24| **Phylogenetically diverse strains are significantly enriched in minimal medium.**

The plot shows log_2 fold change in the relative abundance of bacteria across the four growth conditions. Comparisons in log_2 fold change in relative abundance are performed between minimal shaking microcosm *vs.* complex shaking microcosm and indicated in y-axis, and between minimal standing microcosm *vs.* complex standing microcosm and indicated in x-axis. Circles correspond the isolates and colors indicate phyla. Grey refer to the bacteria that do not significantly change in relative abundance. Stars indicate abundant soil bacteria and hash indicates isolates sharing 100% sequence similarity in the v5v7 regions of the 16s rRNA genes. Several *Actinobacteria* and *Proteobacteria* strains are significantly enriched in minimal medium. In contrast, in complex medium only few bacterial strains are significantly enriched and mainly *Proteobacteria*.

In vivo depletion of highly competitive bacteria strongly alters species richness. Bacteria live in socially entangled multispecies communities where cooperative and competitive interactions can alter the community diversity and structure (Niu *et al.,* 2017). Although counter-intuitive, theoretical studies have reported that competitive interactions promote biodiversity when spatially constrained (Vetsigian *et al.,* 2011, Czárán *et al.* 2001). However, most empirical and theoretical studies employed so far bacterial communities that are very limited in strain number and species diversity. These communities may reflect partially, if at all, natural environments. In order to test the structuring role of the highly competitive strains in a community context, I have *in vivo* depleted the 13 most competitive or the 13 most sensitive strains (defined based on the ABBA screen in Chapter II) from a taxonomically-diverse community composed of 198 members that cover 25 families and four phyla. Perturbation experiments by community members' depletion are performed using exactly the same experimental conditions describe above (complex and minimum medium in shaking or standing) (**Figure 19**). Output communities are profiled *via* amplicons sequencing (V5-V7 regions of the 16S rRNA gene). Importantly, communities that are *in vivo* depleted from the 13 highly competitive (or sensitive) strains are compared to the full community (*i.e.* 198 inoculated strains) from which the same highly competitive (or sensitive) strains are *in silico* depleted (**Figure 20**). The *in silico* depletion is essential to avoid biases and allows a direct and fair comparison between perturbed and unperturbed communities. Each *in vivo* depleted community is compared to a corresponding *in silico* depleted full communities for the same growth medium and state.

The analysis of the alpha-diversity shows that all output communities have lower species richness than the input. Interestingly, *in vivo* depletion of the 13 highly competitive strains leads to a significant and sharp drop in species richness and evenness in almost all growth conditions (**Figure 25**, **panel-A**). Unexpectedly, *in vivo* depleted communities from the 13 highly competitive bacteria in complex medium in shaking show a higher Shannon index than the corresponding *in silico* depleted full communities (**Figure 25**, **panel-A**). The discrepancy between complex medium in shaking and the other conditions is likely caused by the *in silico* depletion of highly competitive isolates. Thus, highly competitive strains contribute mainly in the evenness scores of full communities grown in complex medium in shaking. Taken together, *in vivo* depletion of highly competitive strains considerably reduces species richness in liquid microcosms. By contrast, depletion of the 13 highly sensitive strains marginally affects species richness (**Figure 25**, **panel-A**). Overall, *in vivo* depletion of the 13 highly competitive isolates negatively alters species richness than *in vivo* depletion of the 13 highly sensitive isolates. Although both depletions (*i.e.* highly competitive or highly sensitive) show a similar trend indicating a reduction in the community diversity, *in vivo* depletion of the 13 highly competitive strains has stronger effect on the alpha-diversity than *in vivo* depletion of the 13 highly sensitive strains. These data suggest that highly competitive community members are important community members that promote diversity, which concurs with theoretical-based studies.

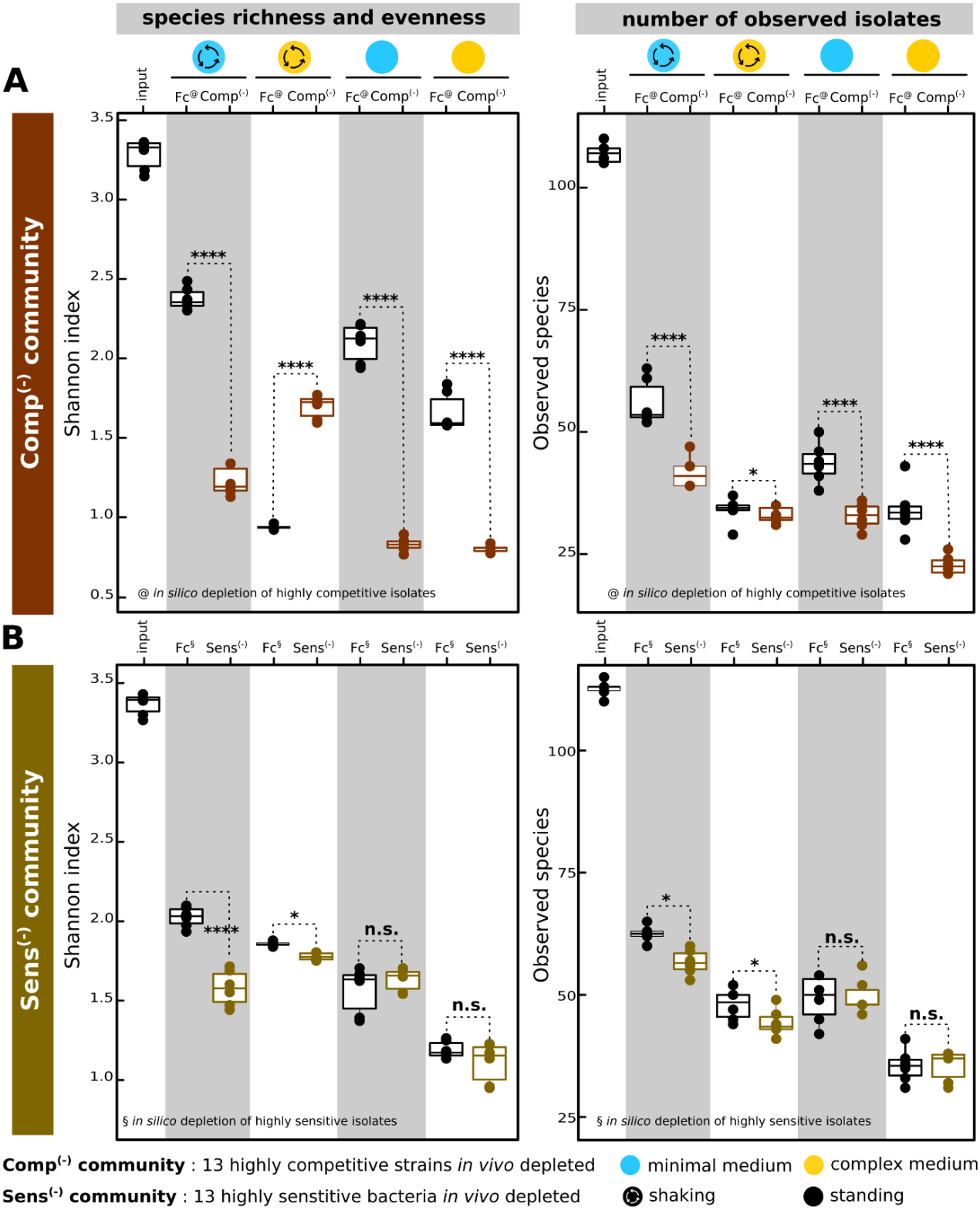


Figure 25| ***In vivo* depletion of the 13 highly competitive bacteria leads to a significant drop in community diversity.**

Panel-A and **-B** depict Shannon index, left-side plot, and Observed species, right-side plot, in Comp$^{(-)}$ communities and Sens$^{(-)}$ communities, respectively. **Panel-A**, *in vivo* depletion of the 13 highly competitive strains leads to a sharp deacrese in species richness. **Panel-B**, *in vivo* depletion of the 13 highly sensitive strains has marginal effect on species richenss.

***In vivo depletion of highly competitive bacteria strongly alters the community structure*.** In order to reveal the effect of both *in vivo* depletions (*i.e.* the 13 highly competitive or sensitive strains) on the bacterial community structure, I calculated Bray Curtis (BC) and weighted UniFrac (wUF) distances within and between the output communities. Contrary to BC distance metric, UniFrac is a beta-diversity measure that incorporate phylogenetic distances between species to compare the samples (Lozupone *et al.*, 2011). In the first instance, I tested whether *in vivo* depletion of the community member alters the homogeneity of the output communities. To this end, I computed the distance to centroid between communities incubated under similar growth conditions, as illustrated in **Figure 26**. The analysis of homogeneity based on BC distances shows no significant changes in the distance to centroid for all output communities. By performing a similar analysis based on wUF instead of BC distances, I could reveal that the distance to centroid for both *in vivo* depletions are significantly higher from the distance to centroid of full community, for all tested microcosms (**Figure 26**, **panel-B**). Interestingly, there are no significant differences in the distance to centroid between *in vivo* depletion of the 13 highly competitive strains and *in vivo* depletion of the 13 highly sensitive strains (**Figure 26**, **panel-B**). The incongruity between BC and wUF distance metrics points to the fact that observed heterogeneity in both *in vivo* depletions compared to the control full community (an increase distance to centroid) has a strong phylogenetic signal that is not measurable by the BC distance metric. Therefore, both depletions primarily alter the phylogenetic homogeneity of the community. Therefore, depletion of the 13 highly competitive or sensitive strains causes a phylogenetic heterogeneity among established communities in the liquid microcosms.

Although the homogeneity test is based a beta-diversity metric, it remains unclear how (dis)similar are *in vivo* depleted communities from the full communities. In the second instance, I quantified the dissimilarity between samples by computing BC and wUF distances between $Comp^{(-)}$ communities (*in vivo* depletion of the 13 highly competitive strains) and full communities *in silico* depleted from highly competitive bacteria and between $Sens^{(-)}$ communities (*in vivo* depletion of the 13 highly sensitive strains) and full communities *in silico* depleted from highly sensitive bacteria. These analyses are depicted as boxplots in **Figure 27** or as principal coordinate analysis in **Supplementary Figure 5**. Interestingly enough, $Comp^{(-)}$ communities are significantly more distant than $Sens^{(-)}$ communities in almost all tested microcosms and according to the two distance metrics (**Figure 27, panel-A** and **-B**). However, an exception is observed in minimal medium in shaking state, $Comp^{(-)}$ communities and $Sens^{(-)}$ communities are equally distant from their respective *in silico* depleted full communities (**Figure 27, panel-B**). These data indicate that *in vivo* depletion of the 13 highly competitive bacteria have more pronounced effect on the community structure than *in vivo* depletion of the 13 highly sensitive bacteria. To further corroborate this analysis, I performed a principal coordinates analysis based on both distance metrics and constrained the analysis by *in vivo* depletion of the 13 highly competitive or sensitive strains (**Figure 28**). According to the BC distance metric, *in vivo* depletion of the 13 highly competitive bacteria

explains up to 7.5% of the observed variance with a *p*-value of 0.001 (**Figure 28, panel-A**).

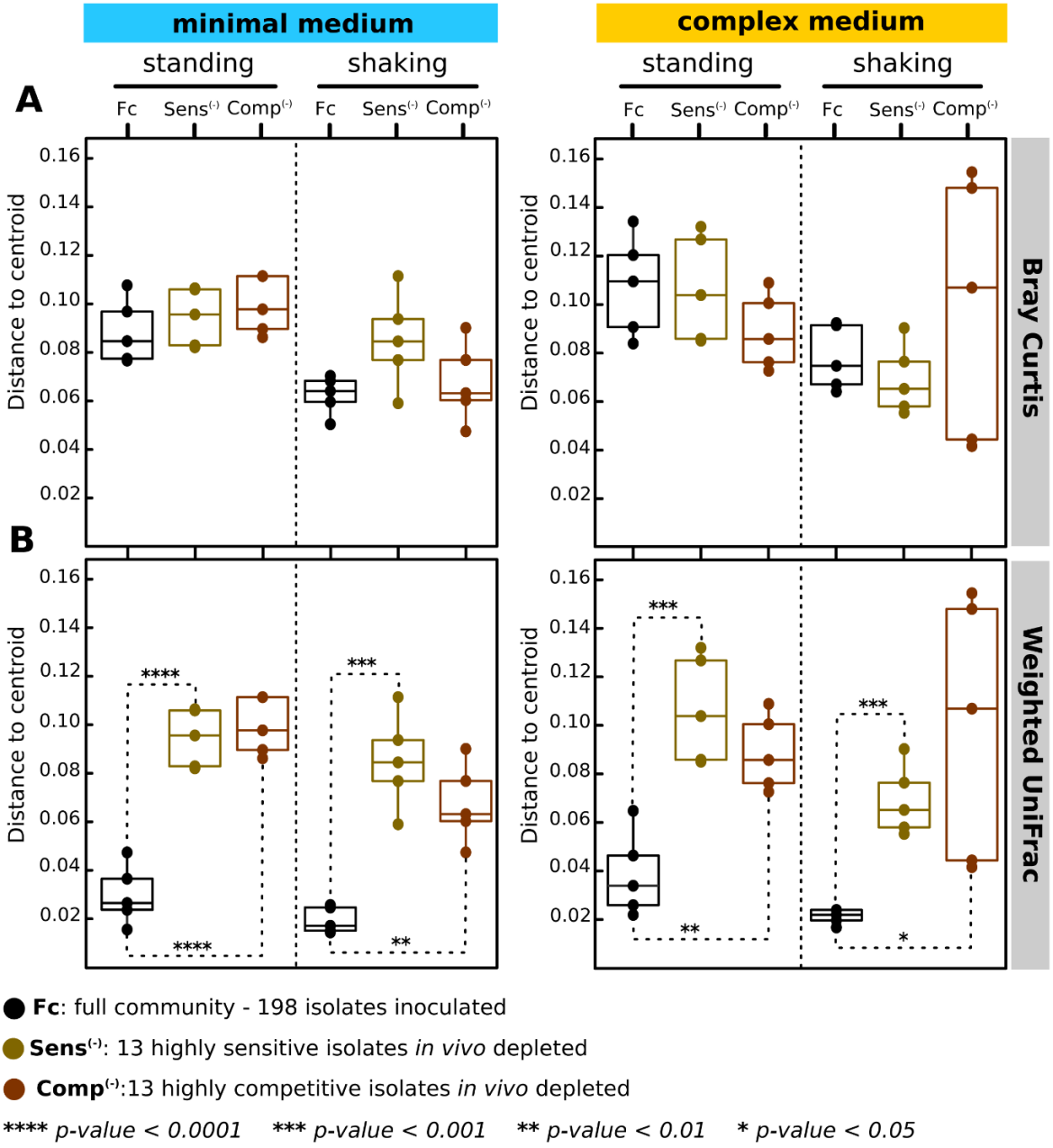


Figure 26| **Both community members' depletions alter the phylogenetic homogeneity of bacterial communities.**

Panel-A and **-B** depict Bray Curtis (BC) and weighted UniFrac (wUF) distances to centroid, respectively. Grouping factor correspond to the four microcosms for each of the three bacterial communities. Each circle refers to a community. **Panel-A**, *in vivo* depletion of the 13 highly competitive or sensitive strains does not significantly alter the community homogeneity, though a trend point toward an increase in distance to centroid in perturbed communities. **Panel-B**, depletion of the 13 highly competitive or sensitive strains leads to a significant increase in the distance to centroid. A higher distance to the centroid indicates that perturbed communities (Comp$^{(-)}$ or Sens$^{(-)}$) are phylogenetically more heterogeneous than full communities under a similar microcosm. Perturbed communities are not significantly different from each other under comparable microcosm.

In counterpart, *in vivo* depletion of the 13 highly sensitive bacteria only explains 5.4% of the variance with a *p-value* of 0.003 (**Figure 28, panel-A**). Similarly, the principal coordinates analysis based on the wUF metric corroborates the above findings, but with lower percentages of explained variance (~3 fold lower) (**Figure 28, panel-B**). These data indicates that both *in vivo* depletions (13 high competitive or 13 highly sensitive) significantly alter output community structure. Nonetheless, *in vivo* depletion of the

13 highly competitive strains alters more strongly bacterial community structure than *in vivo* depletion of the 13 highly sensitive strains. This results indicates that highly competitive community members are not only important for promoting bacterial community diversity, but also for community structure/stability.

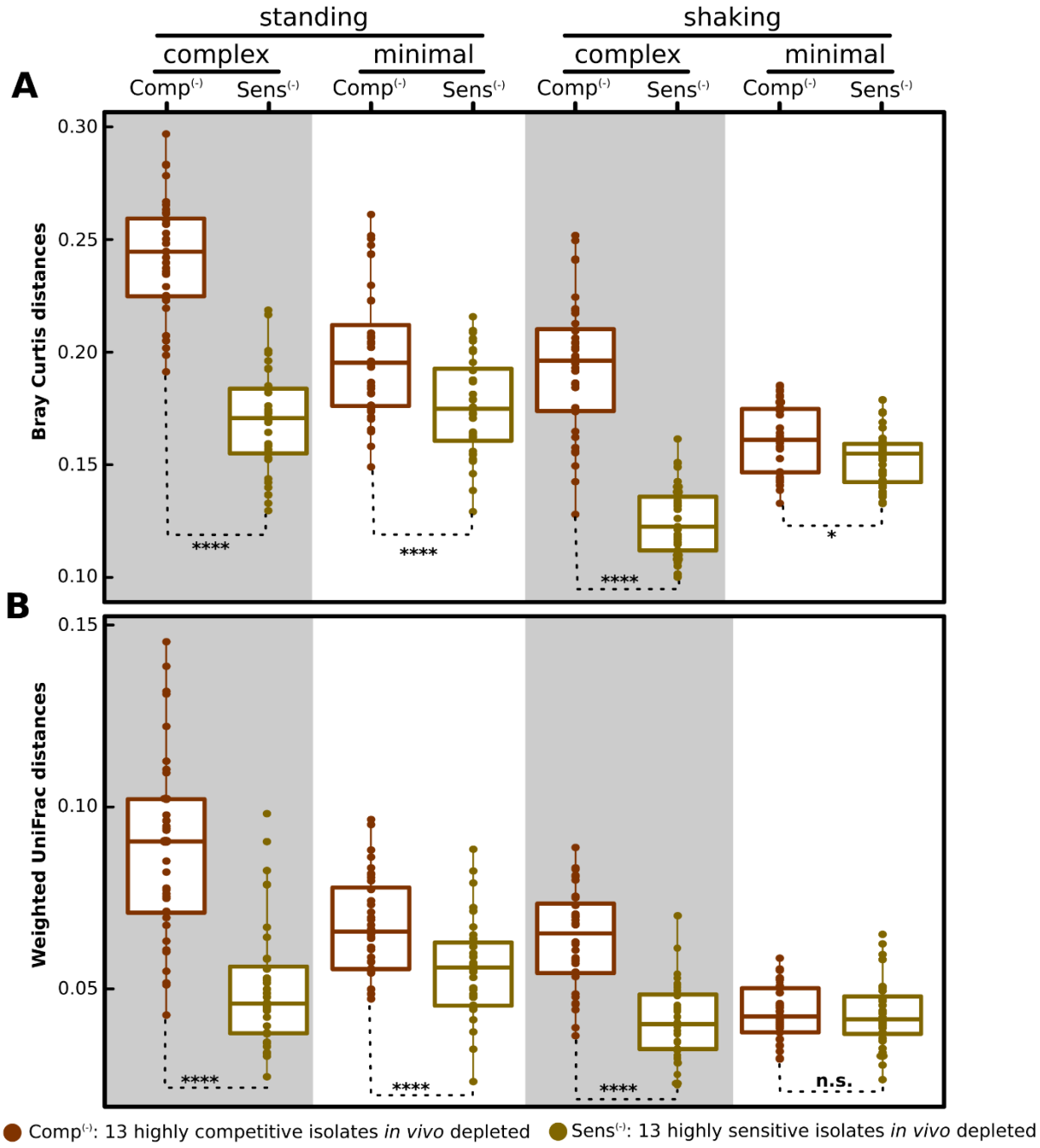


Figure 27| ***In vivo* depletion of 13 highly competitive bacteria alters more strongly community structure.**

Panel-A and **-B** depict Bray Curtis (BC) and weighted UniFrac (wUF) distances between perturbed communities and full community, respectively. Each circle represents a perturbed community and distance of each circle to x-axis translates distance between $Comp^{(-)}$ community and full community *in silico* depleted from highly competitive isolates or between $Sens^{(-)}$ community and full community *in silico* depleted from highly sensitive bacteria. Color code indicates $Comp^{(-)}$ community or $Sens^{(-)}$ community. Stars indicate significance level after Kruskal-Wallis-Conover test, *p-values* are adjusted after Benjamini, Hochberg method. **Panel-A**, $Comp^{(-)}$ communities are significantly more distant from corresponding *in silico* depleted full community than $Sens^{(-)}$ communities are from their corresponding *in silico* depleted full community. **Panel-B**, "wUF" distances show similar pattern as indicate panel-A with the exception that $Comp^{(-)}$ community and $Sens^{(-)}$ community in minimal medium under shaking state are equally distant from correspond *in silico* depleted full community.

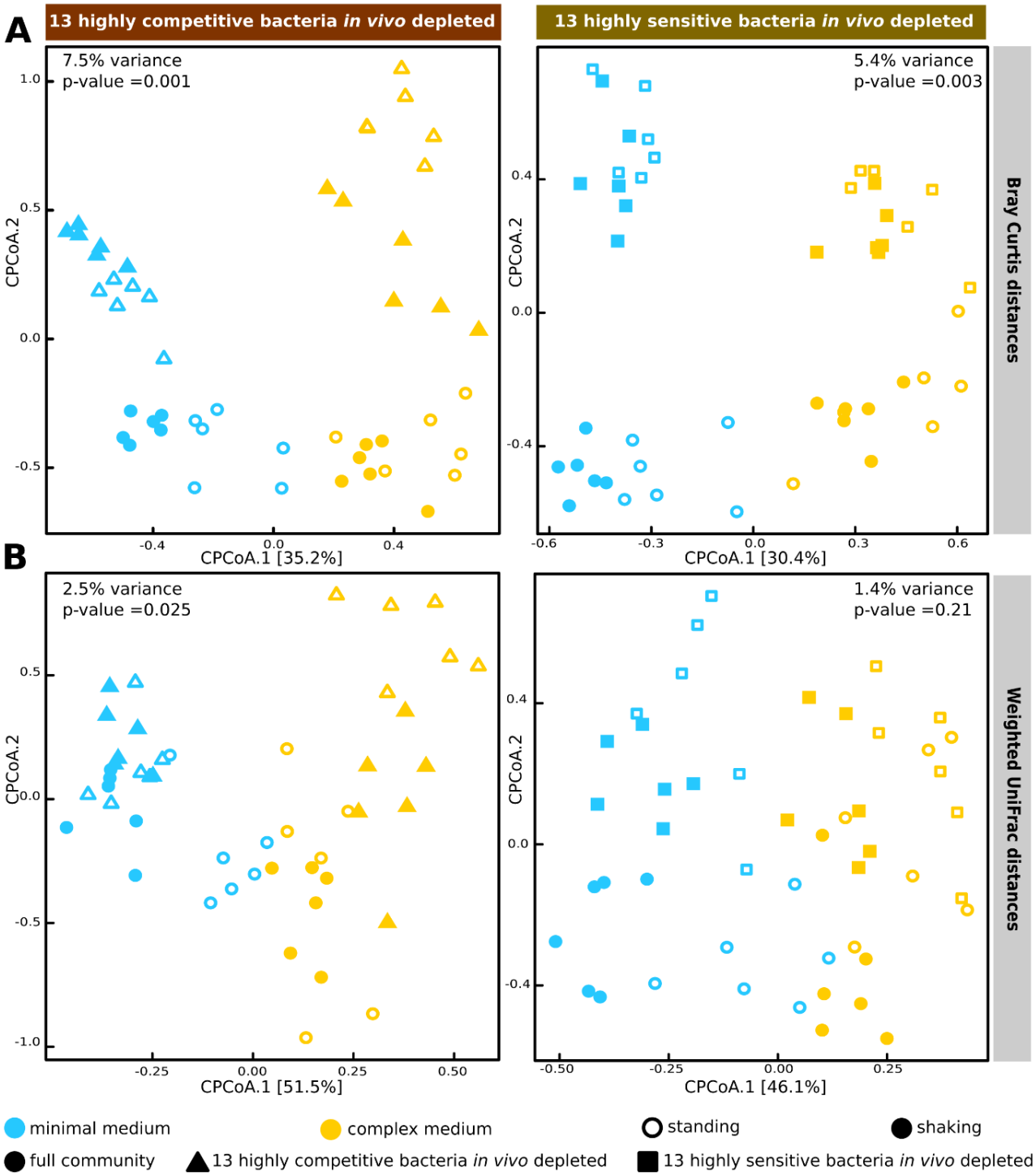


Figure 28| ***In vivo* depletion of the 13 highly competitive strains explains more of variance then *in vivo* depletion of the 13 highly sensitive bacteria.**

Panel-A and **-B** depict constrained principle coordinates analysis of Bray Curtis (BC) and weighted UniFrac (wUF) distances, respectively. PCoA plots are constrained in the first axis by medium and in the second axis by *in vivo* depletion of the 13 highly competitive bacteria in left-side or by *in vivo* depletion of the 13 highly sensitive bacteria in right-side of both panels. **Panel-A**, *in vivo* depletion of the 13 highly competitive bacteria explains 7.5% of the variance and *in vivo* depletion of 13 highly sensitive bacteria explains 5.4% of the variance. **Panel-B**, constrained coordinates analysis on wUF distances show similar pattern as the analysis on BC, but overall explained variance is 3 fold less in wUF analysis then in BC analysis.

To further quantify the shift in the community structure upon *in vivo* depletion of the 13 highly competitive or sensitivity strains within each liquid microcosm, I performed the analysis of similarity (Clarke, 1993) and the permutational multivariate analysis of variance (Anderson, 2001) on BC and wUF distances between Comp$^{(-)}$ or Sens$^{(-)}$ communities and the corresponding *in silico* depleted full communities. Former and latter analyses are respectively represented in **Figure 29** and **Figure 30**. The former analysis compares the mean of ranked dissimilarities between groups (*i.e.* between full communities and Comp$^{(-)}$ or Sens$^{(-)}$ communities under comparable microcosm) to the mean of ranked dissimilarities within groups (*i.e.* between the biological and technical replicates of full communities or Comp$^{(-)}$ communities or Sens$^{(-)}$ communities under a similar microcosm) (**Figure 29**). The analysis of similarity provides a (dis)similarity index that is subjected to a significance test by permutation (999 permutations). An index of value 1 indicates that the two data sets are fully dissimilar, whereas an index of 0 indicates that compared samples are not dissimilar. The threshold of the (dis)similarity index is by consensus fixed between 0.2-0.25 for a significance level of 0.05. The analysis of similarity based on BC distances indicates that the community structure of both perturbed communities are significantly different from the corresponding *in silico* depleted full communities in a given microcosm. One notable exception is observed for Sens$^{(-)}$ communities when grown in a complex medium in standing state ((dis)similarity index <0.25, *p-value* >0.05) (**Figure 29**, **panel-A**). Interestingly, Comp$^{(-)}$ communities show significantly greater (dis)similarity index than Sens$^{(-)}$ communities, except in a minimal medium in standing state (**Figure 29**, **panel-A**). The analysis of similarity based on wUF distances corroborates the finding that removal of the highly competitive community members have stronger effect on the community structure than the removal of the highly sensitive strains (**Figure 29**, **panel-B**). Moreover, Comp$^{(-)}$ communities show significant higher (dis)similarity indexes than Sens$^{(-)}$ communities under all tested microcosms (**Figure 29**, **panel-B**). Taken together, these data suggest that the highly competitive bacterial strains defined based on the ABBA screen maintain bacterial diversity and profoundly shape the community structure in liquid microcosms. Although the analysis of similarity corroborates constrained principal coordinates findings, this analysis is known to be very sensitive to group dispersal and can confound between truly significant differences between groups and within group dispersal (Anderson and Walsh, 2013, Warton *et al.*, 2012). In contrast to the analysis of similarity, permutational multivariate analysis of variance is more robust against group dispersion (Anderson and Walsh, 2013, Warton *et al.,* 2012). To overcome such potential confounding effects, I computed permutational multivariate analysis of variance on BC and wUF between Comp$^{(-)}$ communities or Sens$^{(-)}$ communities and corresponding *in silico* depleted full communities. The multivariate test provides a percentage of explained and residual variance. The significance of explained variance is evaluated by a permutational significance test (999 permutations) (Anderson, 2001). According to BC distances, this test shows that *in vivo* depletion of the 13 highly competitive strains from all liquid microcosms explains significantly more of the observed variance than *in vivo* depletion of the 13 highly sensitive bacteria (**Figure 30**, **panel-A**). In minimal medium in standing condition, both *in vivo* depletions appear to explain, however,

comparable percentage of variance (**Figure 30**, **panel-A**). Interestingly, the analysis of variance of wUF distances clearly shows that *in vivo* depletion of the 13 highly competitive bacteria explains significantly more of the variance than *in vivo* depletion of the 13 highly sensitive bacteria (**Figure 30**, **panel-B**). Taken together, these analyses indicate that both *in vivo* depletion of the 13 highly competitive or sensitive strains alter the community structure. More importantly, the former *in vivo* depletion has a stronger effect on the community structure than the latter. By combining both studies (alpha- and beta-diversity analyses), highly antagonistic community members have an important role in shaping the structure of the community and in promoting the bacterial diversity.

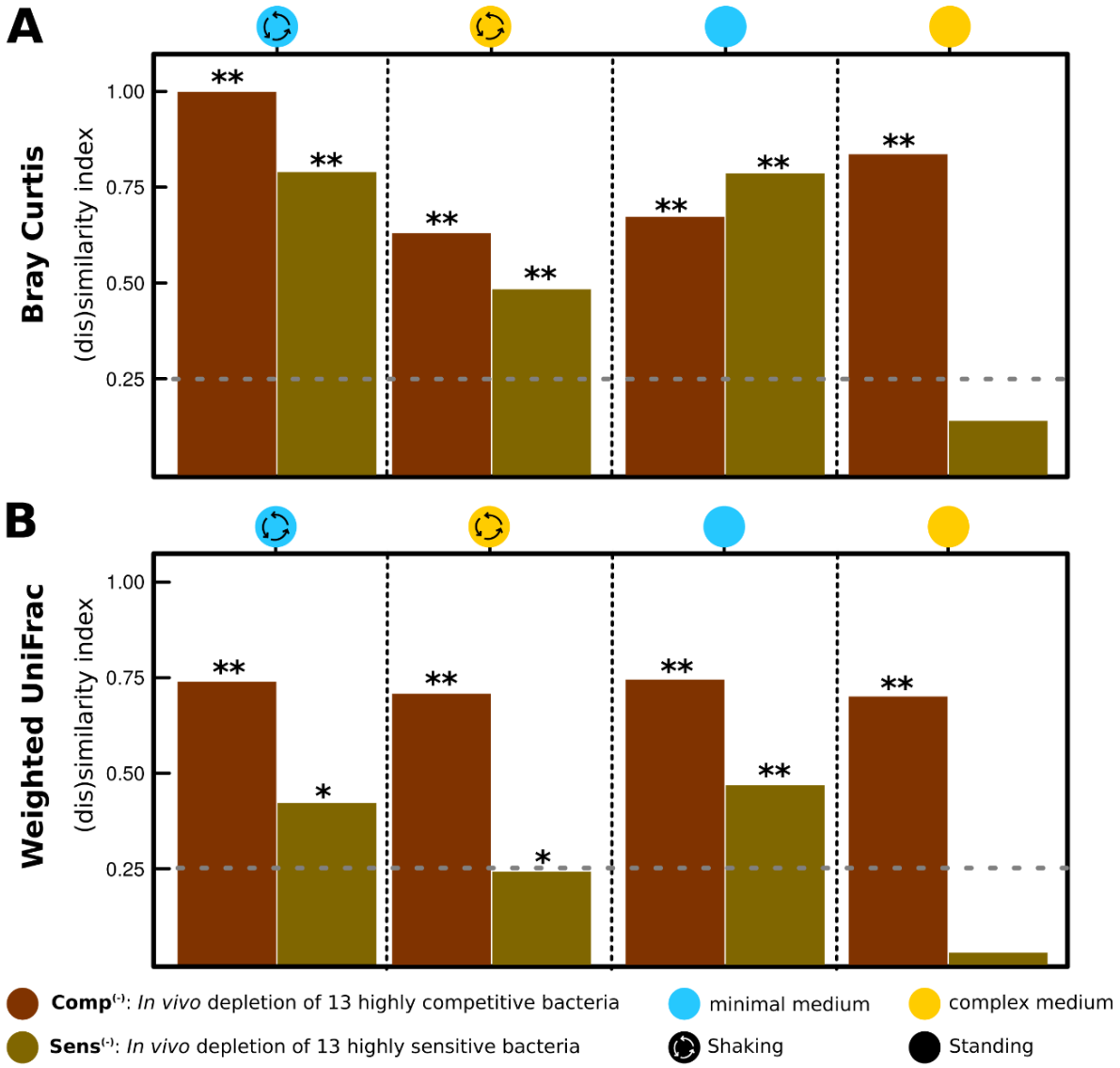


Figure 29| ***In vivo* depletion of 13 highly sensitive bacteria has moderate effect on community structure.**

Panel-A and **-B** depict the analysis of similarity between perturbed communities and full communities computed on Bray Curtis (BC) and weighted UniFrac (wUF) distance metrics, respectively. Analysis of similarity tests for significant differences between the structure of Comp $^{(-)}$ communities and full communities *in silico* depleted from highly competitive bacteria and between Sens$^{(-)}$ communities and full communities *in silico* depleted from highly sensitive bacteria, both analyses are done under comparable microcosm. Threshold for significance in (dis)similarity index is >0.25 for p.value <0.05. **Panel-A**, analysis of similarities based on "BC" distance metric shows that both *in vivo* community members' depletions leads to significantly different communities, except *in vivo* depletion of 13 highly sensitive bacteria in complex medium under standing state. The analysis reveals that almost all Comp$^{(-)}$ communities show higher (dis)similarity index than Sens$^{(-)}$ communities, except in minimal medium under standing state. **Panel-B**, under all growth conditions Comp$^{(-)}$ communities show significantly higher (dis)similarity index than Sens$^{(-)}$ communities.

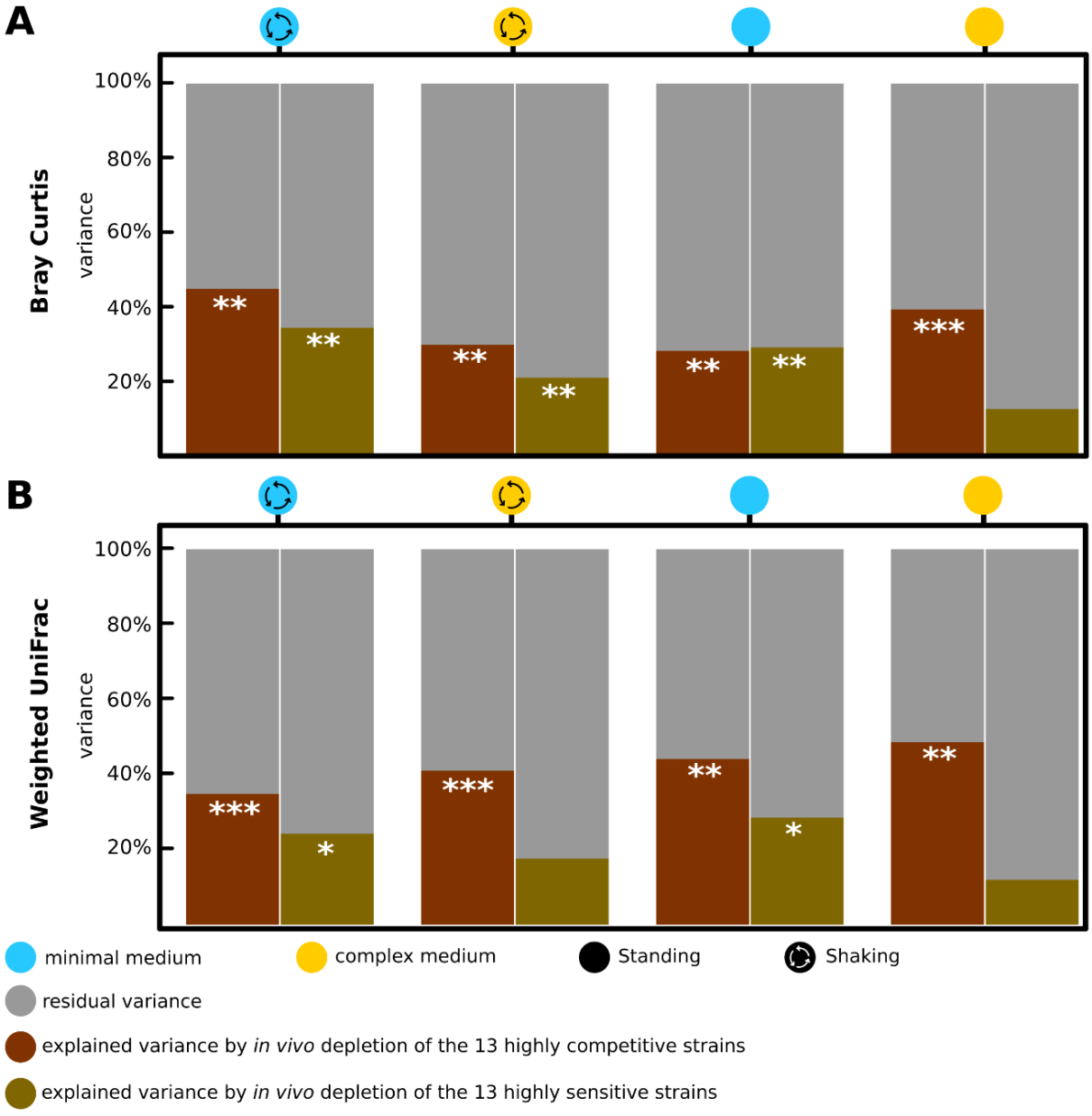


Figure 30| ***In vivo* depletion of 13 highly competitive bacteria alters more strongly community structure than *in vivo* depletion of 13 highly sensitive bacteria.**

Panel-A and **-B** show permutational multivariate analysis of variance of Bray Curtis (BC) and weighted UniFrac (wUF) distances between perturbed communities and full communities, respectively. Color code in stack bar indicates residual variance, variance explained by *in vivo* depletion of the 13 highly competitive strains or the 13 highly sensitive bacteria. 999 permutations were used to determine significance level of the explained variance. **Panel-A**, *in vivo* depletion of the 13 highly competitive strains explains more variance than *in vivo* depletion of the 13 highly sensitive bacteria, except in minimal medium under standing state where both community perturbations show comparable percentage of explained variance. **Panel-B**, *in vivo* depletion of the 13 highly competitive strains explains more variance than *in vivo* depletion of the 13 highly sensitive strains under all tested microcosms.

Several Proteobacteria strains significantly increase in the relative abundance upon removal of highly competitive community members. In order to explain the shift in community structure after *in vivo* depletion of the 13 highly competitive or sensitive strains, I calculated $\log_2$ fold-change in the relative abundance of the bacteria after both *in vivo* depletions in each microcosm. The relative abundance of the isolates can either significantly increase or decrease, or do not significantly change in $Comp^{(-)}$ communities or $Sens^{(-)}$ communities compared to the full *in silico* depleted communities. The raw data used for the analysis of the enrichment test are presented in **Supplementary Figures 6 and 7** and a synthesis of these analyses is illustrated in **Supplementary Figure 8** and **Figure 31**. By comparing $Sens^{(-)}$ communities with full communities, I identified 26 isolates that show a significant decrease in their relative abundance and only 5 isolates that show the opposite trend (**Figure 31**, **panel-A** and **supplementary Figure 8 , panel-A**). Remarkably, the bacterial isolates n°60 (*Microbacteriaceae* family) reproducibly show a decrease in its relative abundance in all tested microcosms (**Figure 31**, **panel-A**). Notably, five isolates (n°63, 9, 1280, 627 and 170) grow better in the absence of sensitive strains, but only in minimal medium. This result suggests that the removal of sensitive bacteria do not provide a competitive advantage for the other community members (**Figure 31**, **panel-A**). These data indicate that observed shift in the community structure after *in vivo* depletion of the 13 highly sensitive bacteria is mainly explained by a significant decrease in the relative abundance of several strains in perturbed communities. Based on the fact that the growth of many isolates is impaired in the absence of the sensitive strains, it is plausible that highly sensitive bacteria engage in cooperative interactions with community members that show altered growth after their removal. This could represent an alternative way to persist within a complex community without the need to compete through the secretion of antimicrobials.

In contrast to *in vivo* depletion of the 13 highly sensitive bacteria, *in vivo* depletion of the 13 highly competitive bacteria leads to significant growth increase of 30 isolates and growth decrease of 14 isolates (**Figure 31**, **panel-B** and **supplementary Figure 8, panel-B**). Remarkably, the relative abundance of the isolate n°9, a *Pseudomonas* strains, significantly increases by several order of $\log_2$ fold in the absence of the competitor strains in all tested microcosms (**Figure 31**, **panel-B**). Remarkably, the relative growth of several β-*Proteobacteria* (isolates, n°170, 83, 70, 411, 318D1, 219 and 267), significantly increases in complex medium under both states, whereas the relative abundance of several *Actinobacteria* isolates increases in minimal medium (**Figure 31**, **panel-B**). *In vivo* depletion of the 13 highly competitive bacteria primarily leads to the enrichment of *Proteobacteria* isolates and to an over dominance of the *Pseudomonas* strains n°9. It is clear from this analysis that *in vivo* depletion of the 13 highly competitive or sensitive bacteria alter the bacterial community structure differently, with sensitive bacteria primarily promoting the growth of specific community members and competitive bacteria broadly restricting the growth of phylogenetically diverse taxa. Overall, both *in vivo* depletions alter community structure, but similarly. Although *in vivo* depletion of the 13 highly sensitive bacteria alters community composition and structure, *in vivo* depletion of the 13 highly competitive bacteria has

a stronger effect on species richness and community structure in all tested conditions. More importantly, *in vivo* depletion of the highly competitor strains leads to enrichment of mainly *Proteobacteria* isolates and more particularly to the over-dominance of a *Pseudomonas* strain. Our results therefore suggest that the secretion of antimicrobials could prevent overgrowth of numerous isolates, therefore increasing community diversity and stability. These data corroborate the general assumption that competitive interactions between species are pillar foundations for biodiversity and community stability (Stubbendieck *et al.,* 2016). However, to what extent this community behaviors can be extrapolated to *in planta* experimental system is explored in the following.

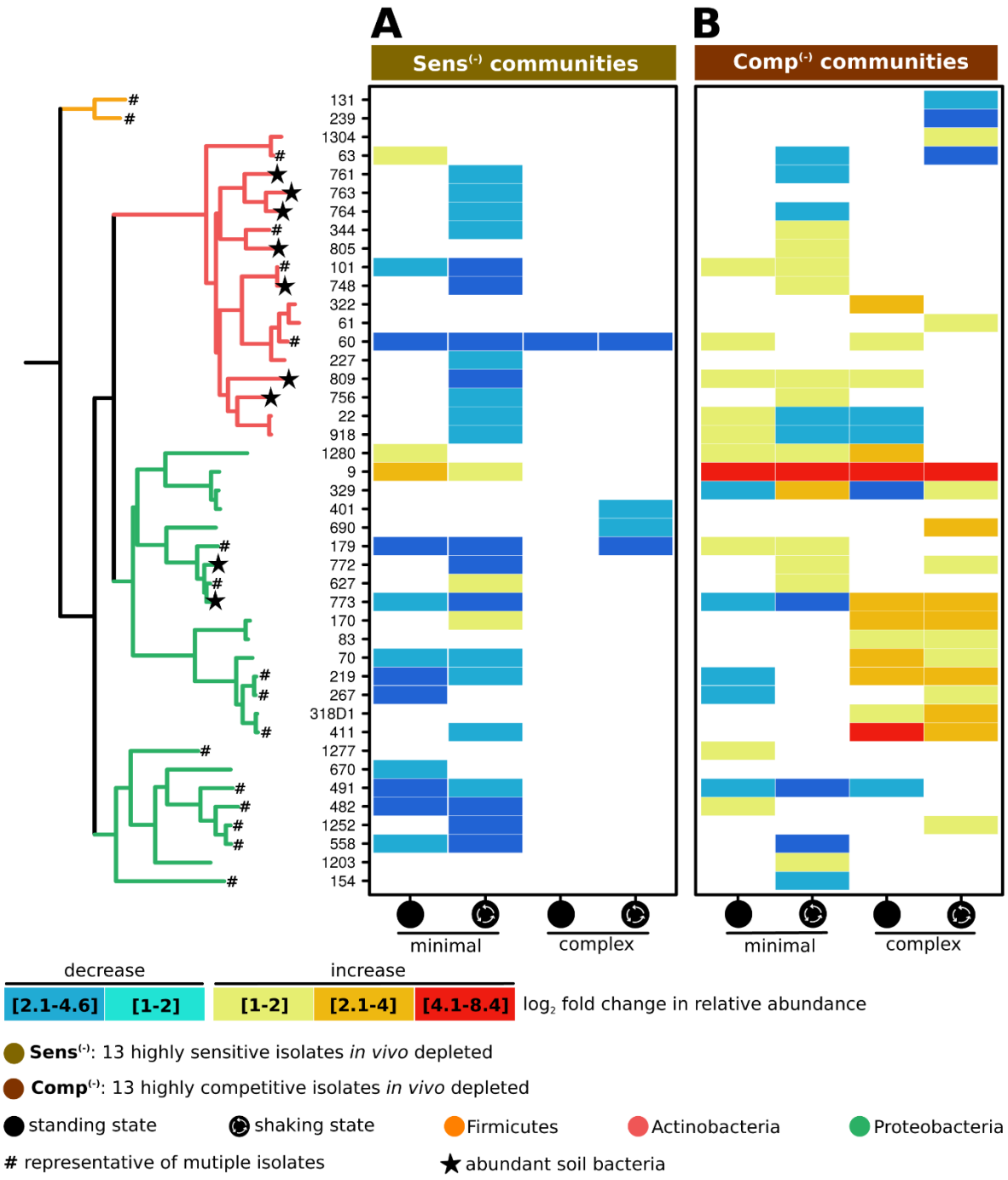


Figure 31| ***In vivo* depletion of 13 highly competitive bacteria causes a dramatic shift in community structure.**

The heatmap plots depict the isolates that significantly increase or decrease in relative abundance upon perturbation by community members' depletion. Bacterial isolates are indicated in y-axis. Star indicates abundant soil bacteria and hash indicates representative isolates. The isolates are ordered by phylogeny of the v5v7 regions of the 16s rRNA genes. Color code in the tree indicates phylum. Significant decrease in relative abundance is color coded by a blue gradient and significant increase in relative abundance is color coded by a red gradient in the both heatmaps. **Panel-A**, *in vivo* depletion of the 13 highly sensitive strains leads mainly to a significant decrease in the relative abundance of several bacterial isolates. **Panel-B,** *in vivo* depletion of the 13 highly competitive strains leads to a significant increase in the relative abundance of several *Proteobacteria* and *Actinobacteria* isolates. The *Pseudomonas* isolates n°9 highly increases in relative abundance across all perturbed microcosms upon depletion of the 13 highly competitve strains.

III.C. The *A. thaliana* root bacterial microbiota remains resilient against the applied perturbations

Assemblage of distinct bacterial communities in the roots and the matrix of calcined clay and calcined clay amended with 3% peat. Land plants use soil as support to grow and uptake minerals. However, natural soils cannot be employed in gnotobiotic system since it is not sterile, it is autoclave-mediated sterilization alters the soil chemistry. To best mimic natural soil conditions for plant growth, I used the inorganic matrix calcined clay to grow *Arabidopsis thaliana* in a closed environment under controlled laboratory conditions. Calcined clay experimental system has been already employed in several studies to understand microbiota establishment in the shoot and root of *Arabidopsis thaliana* (Bai *et al.*, 2015, Lebeis *et al.*, 2015). In order to study the role inter-bacterial interactions in the establishment of the *A. thaliana* root-associated microbiota, I employed two different gnotobiotic experimental systems; 1- calcined clay, 2- calcined clay amended with 3% peat (**Figure 32**). Peat in the latter experimental system substitutes complex organic matter that can be found in natural soil that are depleted from the calcined clay. Seeds were sowed in the matrix of both gnotobiotic systems that have been per-inoculated with three different synthetic bacterial community; 1- full community (198 bacterial isolates and the same set of isolates used in liquid microcosms), 2- $Comp^{(-)}$ (*i.e.* the community is depleted from the 13 highly competitive strains) or 3- $Sens^{(-)}$ (*i.e.* the community is depleted from the 13 highly sensitive strains). The experimental design for bacterial communities employed for *in planta perturbations* is similar to the one employed for liquid microcosms (**Figure 20**). After 7 weeks of incubation in phytochambers, the matrix (*i.e.* bulk clay) and the roots (*i.e.* roots from three to four plants pooled together and washed from clay particles) have been harvested and the bacterial communities were profiled by sequencing the v5v7 regions of the 16s rRNA genes. The shoots (above-ground plant tissue) have been freshly weighted (further details provided in Materials and Methods). Interestingly, neither the peat treatment nor the depletion of community members alter significantly the shoot fresh weight (**Supplementary Figure 10)**. The above-ground biomass is not affected by the treatments or by the community perturbations.

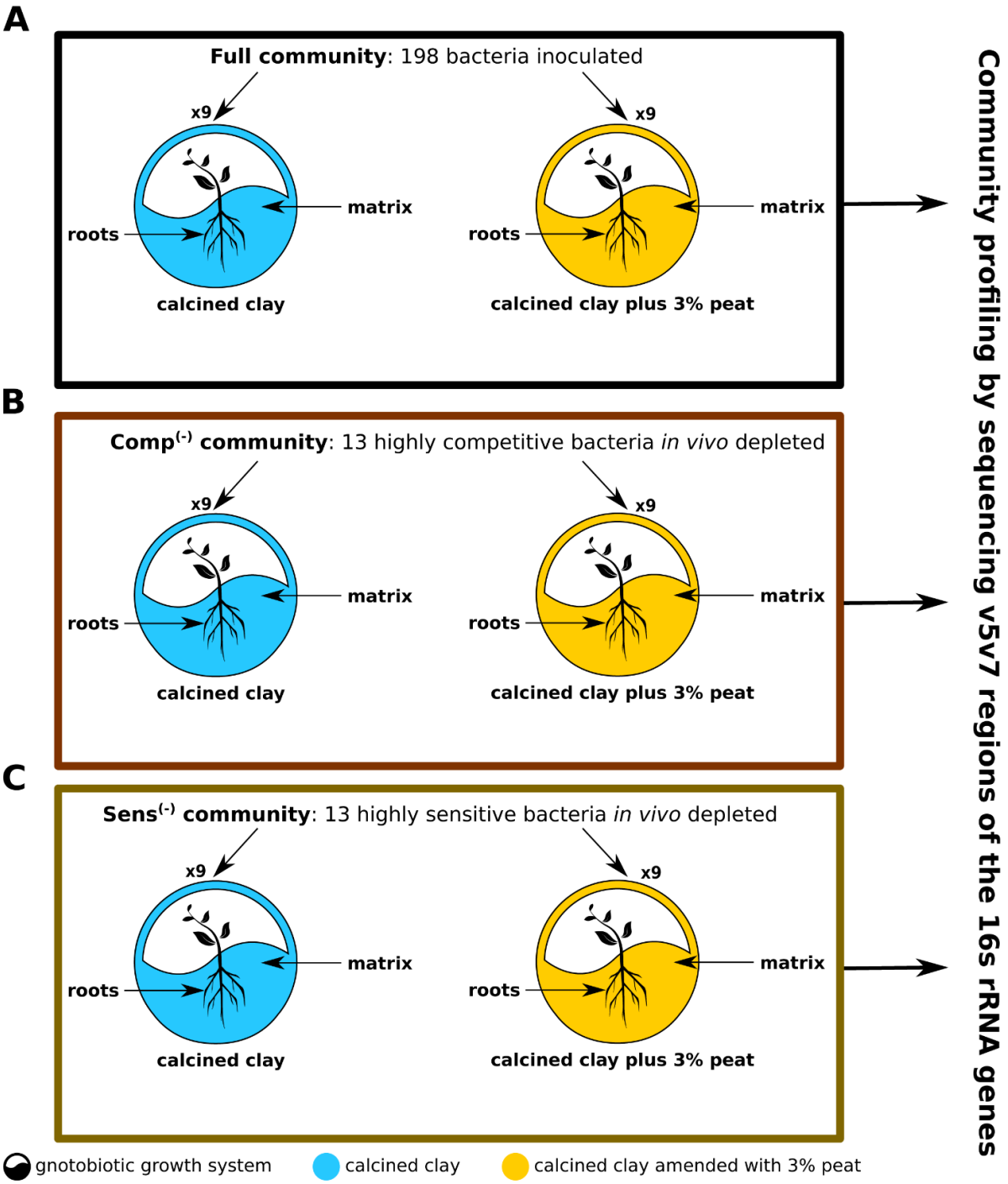


Figure 32| **Experimental setup of *in planta* perturbation by community members' depletion experiment.**

Flowcharts depicting experimental design of perturbation by community members' depletion experiment. Two different growth conditions are included and indicated by color code; 1- matrix depleted from organic (100g of autoclaved calcined clay). 2- matrix amended with complex organic matter (100g of autoclaved calcined clay plus 3% autoclaved peat). *Arapidopsis thaliana*, Col-0, were grown in both gnotobiotic conditions for 7 weeks. Each growth condition include 9 pots and each pot contain 4-5 plants. **Panel-A**, 198 bacterial isolates are inoculated into the two growth conditions, this condition is referred as full community. **Panel-B**, the 13 highly competitive strains are *in vivo* depleted from the full community and indicated by Comp$^{(-)}$ community. **Panel-C**, the 13 highly sensitive strains are *in vivo* depleted from full community and indicated by Sens$^{(-)}$ community.

In the first instance, I sought to evaluate whether distinct bacterial communities assemble in the roots and the matrix of both tested gnotobiotic systems, calcined clay and calcined clay plus 3% peat. To further reveal whether 3% peat amendment affects the bacterial community diversity, I computed two alpha-diversity measures (Shannon index and Observed species) in the input and the output communities (**Figure 33**). This analysis reveals that all output communities display a significantly lower Shannon index and Observed species score than input communities (**Figure 33, panel-A and -B**). Interestingly, the matrix of calcined clay shows a significant low Shannon index than the roots in the same system (**Figure 33, panel-A**). However, more bacterial species are observed in the matrix than in the roots in calcined clay (**Figure 33, panel-B**). In contrast, the matrix and the roots of the calcined clay plus 3% peat system show comparable Shannon index (**Figure 33, panel-A**) and Observed species scores (**Figure 33, panel-B**). These data indicate that complex organic matter (peat or *A. thaliana* roots exudation) causes a reduction in species richness of bacterial communities in gnotobiotic system. These findings are supported by the liquid microcosm observations that indicated that higher species richness is maintained in minimal medium. Although the matrix and the roots in the calcined clay plus 3% peat system show comparable alpha-diversity scores, these observations do not exclude that the bacterial communities in former and latter habitat have a similar community structure.

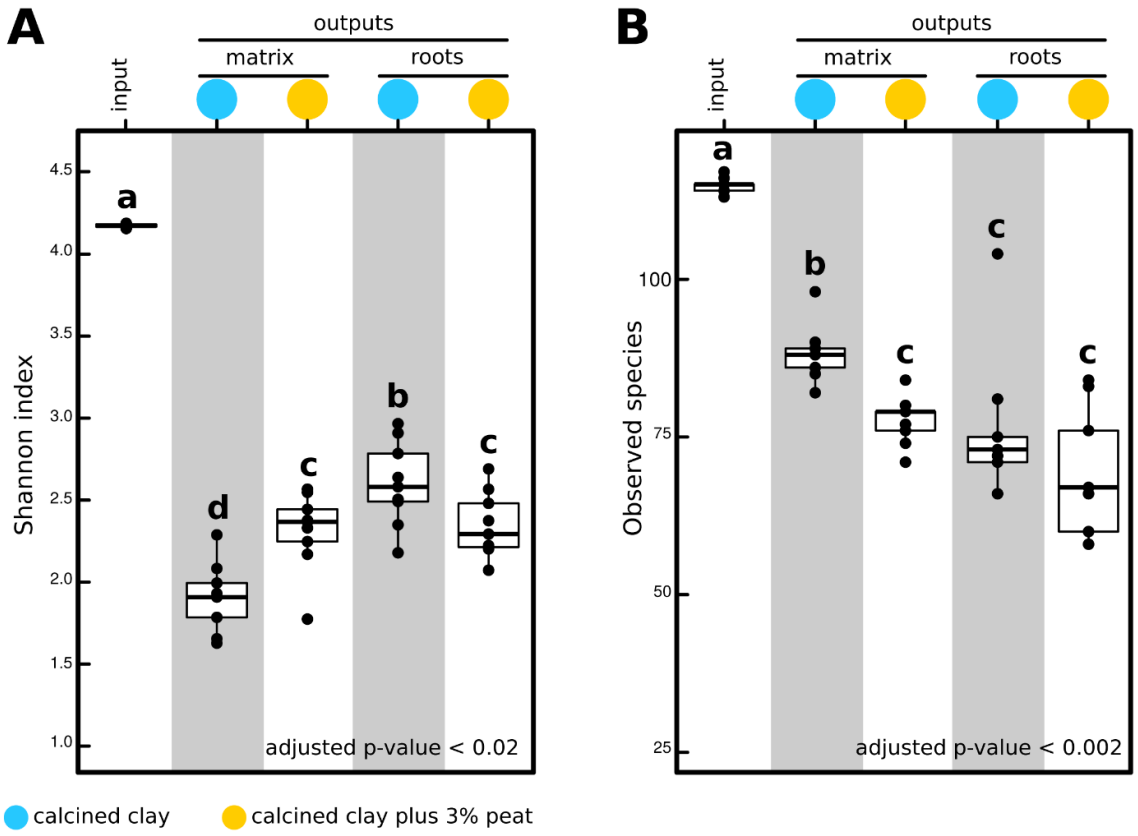


Figure 33| **A higher bacterial diversity is observed in the matrix of calcined clay.**

Box-plots depicting scores in alpha-diversity of input and output communities. **Panel-A**, Shannon index measures species richness and evenness of the bacteria. The root-associated bacterial communities show significantly higher Shannon index than matrix-associated communities in the calcined clay growth system. **Panel-B**, the matrix of calcined clay maintains a higher species diversity than remaining output communities of roots or the matrix of calcined clay plus 3% peat.

To further reveal the dissimilarities in the community structure of output communities, I computed Bray Curtis (BC) distances between output communities and applied principal coordinate analysis on these distances (**Figure 34, panel-A**). Also, I constrained the coordinate analysis by habitat (matrix or roots) in the first axis and by treatment (3% peat amendment) in the second axis (**Figure 34, panel-B**). Principal coordinates analysis indicates that the habitat (matrix or roots) explains most of the observed variance (30.9%, **Figure 34, panel-A and B**) and the treatment explains up to 13.4% of the observed variance in the community structure of output communities (**Figure-34**, **panel-B**). Furthermore, the analysis of group homogeneity based on BC distances indicates that the matrix communities are more homogeneous than the roots communities (**supplementary Figure 10 panel-B)**. Collectively, the combined beta-diversity analyses indicate that assembled bacterial communities in the matrix and the roots for both systems (*i.e.* calcined clay and in calcined clay plus 3% peat) are distinct by their community structure.

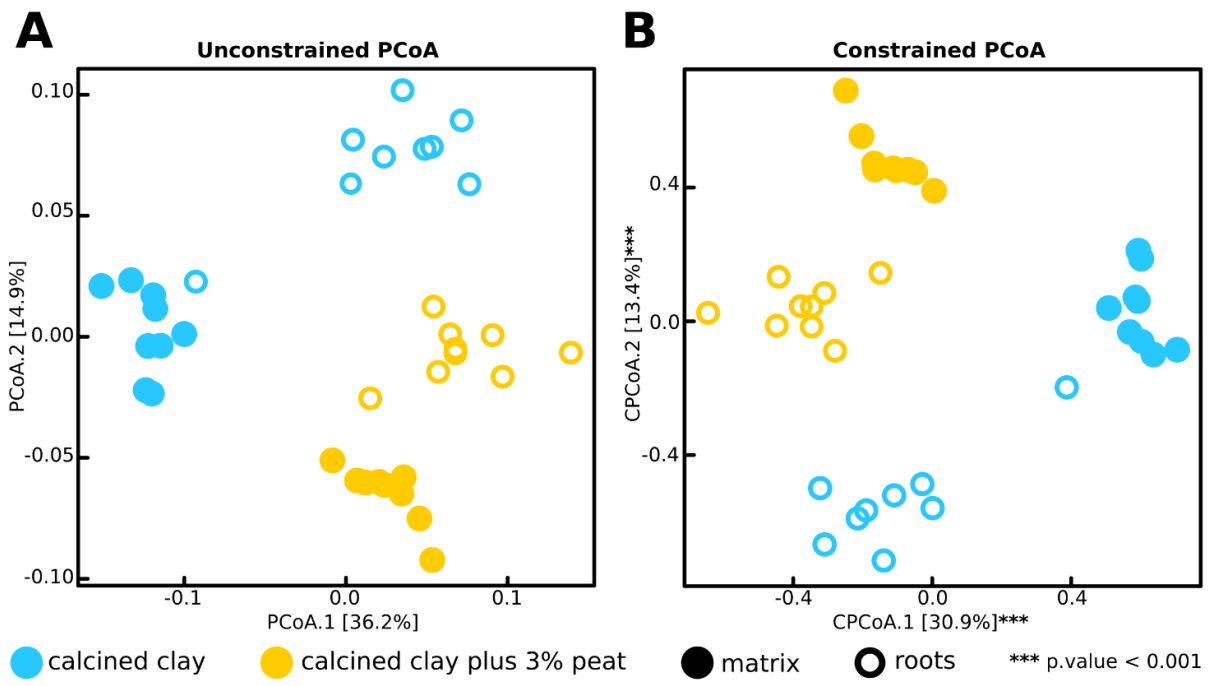


Figure 34| **Assemblage of distinct bacterial communities in matrix and roots in both growth conditions.**

Panel-A and **B** depict unconstrained and constrained Principal Coordinates Analysis (PCoA) on Bray Curtis (BC) distances between matrix and roots samples from calcined clay and calcined clay plus 3% peat conditions, respectively. Each shape represents a bacterial community (one sample). Color code indicates calcined clay or calcined clay plus 3% peat conditions. Filled and unfilled shapes indicate matrix and roots samples, respectively. Distances between samples correspond to (BC) dissimilarity distance computed on normalized count data. **Panel-A**, first axis in PCoA plot explains most of the observed variance in the data and separates between calcined clay matrix samples and calcined clay plus 3% peat matrix and roots samples. **Panel-B**, constrained analysis indicates that 13.4% of the variance is explained by the amendement of 3% peat.

In order to further identify the bacterial isolates that best explain the observed shift in the community structure, I performed an enrichment test for $\log_2$ fold change in the relative abundance. More precisely, I compared the $\log_2$ fold change in the relative abundance of the isolates of calcined clay matrix *versus* calcined clay plus 3% peat matrix and the roots in calcined clay *versus* the roots in calcined clay plus

3% peat (**Figure 35**). Remarkably, several *Actinobacteria* isolates are significantly enriched in calcined clay matrix compared to matrix with peat. (**Figure 35**, y-axis), where more *Proteobacteria* isolates are significantly enriched. On the other hand, several *Proteobacteria* isolates are significantly enriched in the roots in calcined clay system compared to the roots in calcined clay plus 3% peat (**Figure 35**, x-axis). These data indicate that more bacterial isolates are significantly enriched in the matrix and the roots of calcined clay system. Altogether, the analysis of output communities in the matrix and the roots for both systems indicates that the amendment of organic matter causes a reduction in alpha-diversity and alters the community structure in the roots and matrix. Furthermore, the data indicate that *Proteobacteria* isolates grow better in carbon-rich habitats. Interestingly, these striking alterations in the bacterial community composition have no phenotypic effect on the shoot biomass. The tested growth conditions (calcined clay and calcined clay plus 3% peat systems) coupled with the habitats (matrix and roots) lead to the assembly of distinct bacterial communities, which validate the prominent role of the soil edaphic characteristics (*i.e.* carbon content) in altering the structure of bacterial communities.

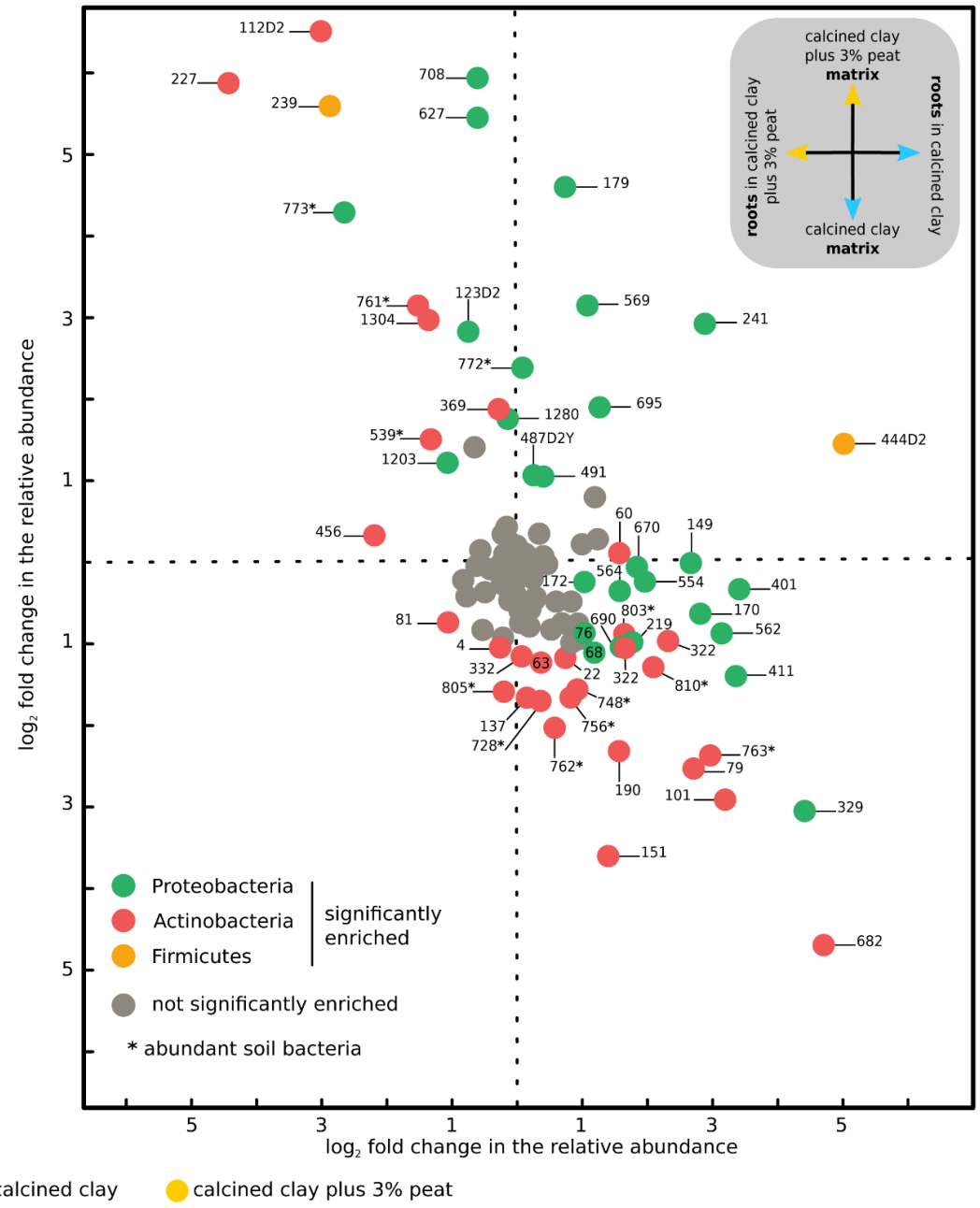


Figure 35| **Several Actinobacteria isolates are significantly enriched in calcined clay matrix.**

The plot shows log_2 fold change in the relative abundance of bacterial isolates. Y-axis depicts the comparison between calcined clay matrix and calcined clay plus 3% peat matrix. X-axis depicts the comparison between roots in calcined clay and roots in calcined clay plus 3% peat of log_2 fold change in relative abundance of isolates. Circles correspond to isolates and colors indicate phyla. Grey color indicates isolates that are not significantly enriched. Star indicates abundant soil bacteria. Several *Actinobacteria* isolates are significantly enriched in calcined clay matrix compared to calcined clay plus 3% peat matrix.

The community diversity remains stable against perturbation by community members' depletions. The analysis of alpha-diversity in output communities from the liquid microcosms indicated that *in vivo* depletion of the 13 highly competitive strains alters negatively the community diversity. In contrast, *in vivo* depletion of the 13 highly sensitive strains had minor to no effect on the community diversity. In order to further reveal whether similar perturbations also lead to a similar phenotypes in the matrix and the roots, I *in vivo* depleted the same 13 highly competitive or sensitive strains from the full community members (refer to liquid microcosms). The analysis of the Shannon index and the Observed species of output communities in the matrix and the roots indicates that *in vivo* depletion of the 13 highly competitive or sensitive strains do not significantly alter the community diversity in both systems (*i.e.* calcined clay or calcined clay plus 3% peat) (**Figure 36, panel-A** and **-B**). However, an exception is reported for the roots in calcined clay plus 3% peat, *in vivo* depletion of the 13 highly competitive strains alters significantly species evenness (**Figure 36, panel-A**) but not species richness (**Figure 36, panel-B**). The analysis of the alpha-diversity metrics of the output communities in both gnotobiotic experimental systems are incongruent with liquid microcosms and rather indicate that the community diversity is maintained upon perturbations in the matrix and the roots. The reported discrepancy between liquid microcosms and gnotobiotic experimental systems is not surprising and may rely on the fact that former and latter experimental systems are fundamental different by their physical and chemical properties. Indeed, liquid microcosms are aquaponic cultures where secreted molecules easily diffuse in the system. In contrast, the diffusion of secreted molecules by bacteria in the matrix are rather constrained by clay moisture. Therefore, it is plausible that secreted antimicrobials have a limited or localized effect (at the micro-scale) on the bacterial diversity that cannot be revealed by amplicons sequencing. Alternatively, the host plant can directly or indirectly fine-tune community members' abundance. Both former and latter hypotheses require further empirical work.

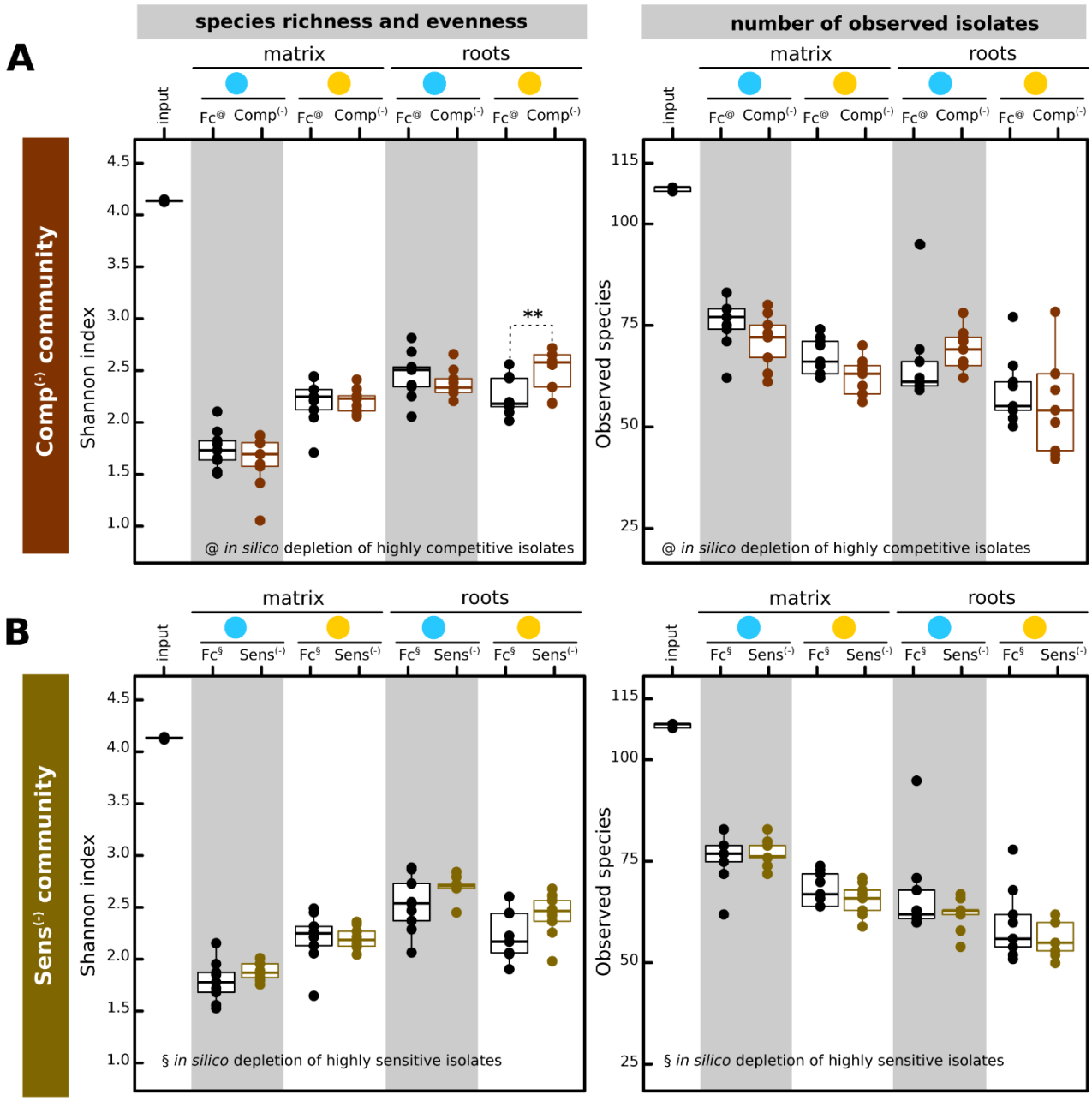


Figure 36| **The bacterial community diversity remains stable against perturbations.**

Panel-A and **-B** depict alpha-diversity measures in Comp$^{(-)}$ communities and Sens$^{(-)}$ communities, respectively. **Panel-A**, *in vivo* depletion of the 13 highly competitive strains do not alter significantly the species richness of bacterial communities in matrix and roots under both conditions. **Panel-B**, *in vivo* depletion of the 13 highly sensitive strains also do not significantly alter the species richness in all output communities.

In vivo community members' depletion alters weakly the structure of the Arabidopsis root-associated bacterial microbiota. In order to study the structure of the bacterial communities *in plant* and in the surrounding matrix, I computed Bray Curtis (BC) and weighted UniFrac (wUF) distances between samples. The analysis of BC and wUF distances between samples sharing the same grouping factor (*i.e.* distances between technical and biological replicates of the matrix or the roots in calcined clay or in calcined clay plus 3% peat) indicates that *in vivo* depletion of the 13 highly competitive or sensitive strains do not alter significantly group homogeneity (**Supplementary Figure 11 panel-A** and **panel-B)**. However**,** depletion of the 13 highly sensitive strains in the calcined clay matrix lead to a significant decrease in the distance to centroid indicating therefore an increase communities' homogeneity (**Supplementary Figure 11 panel-A**). Additionally, the analysis of group homogeneity indicates that root-associated bacterial communities remains unchallenged by applied perturbations (**Supplementary Figure 11 panel-B**). To further reveal whether any of the perturbations induce a shift in the structure of the output bacterial communities, I performed a principal coordinate analysis on BC and wUF distances (**Supplementary Figure 12 panel-A** and **panel-B**) and also constrained this analysis by the applied perturbations (*in vivo* depletion of the 13 highly competitive or highly sensitive strains) (**Supplementary Figure 13 panel-A** and **panel-B**). It is remarkable to report that both *in vivo* depletions explain approximatively a comparable magnitude of observed variance that hardly exceed 5% according to BC distances (**Supplementary Figure 13 panel-A**). In counterpart, constrained PCoA based on wUF distances indicates that *in vivo* depletion of the 13 highly competitive strains slightly alters the bacterial community structure (3.8%), whereas no significant difference is observed upon *in vivo* depletion of the 13 highly sensitive strains (**Supplementary Figure 13 panel-B**). Taken together, these data nonetheless indicate that both perturbations have a very limited effect on the structure of the output communities in the roots and the matrix.

To further quantify the community shift upon perturbation, I computed the analysis of similarity (**Figure 37**) and the permutational multivariate analysis of variance (**Figure 38**) using BC and wUF distance metrics. The analysis of similarity based on BC distances shows that perturbed bacterial communities in the matrix have a higher dis(similarity) score than in the roots (**Figure 37, panel-A**). However, both *in vivo* depletions show roughly similar (dis)similarity indexes which indicates that the depletion *per se*, rather than the competitiveness potential, significantly alters the community structure (**Figure 37, panel-A**). Alternatively to BC distances, the same analysis based on wUF distances indicates that only *in vivo* depletion of the 13 highly competitive strains alters significantly the community structure in the matrix (**Figure 37, panel-B**). Interestingly, the analysis of *in vivo* depletion of the 13 highly sensitive strains from the roots in calcined clay shows a significant (dis)similarity index above 0.25 (**Figure 37, panel-B**). Collectively, the analysis of similarity indicates that perturbed communities in the roots are weakly dissimilar from unperturbed communities, whereas perturbed communities in the matrix are more dissimilar from unperturbed communities. Furthermore, these findings are corroborated by the multivariate analysis of variance. The analysis indicates that *in vivo* community members' depletions

explain significantly less variance in the roots than in the matrix (**Figure 38**). Unlike the variance calculated based on BC distances that show no difference between the two perturbed communities (**Figure 38A**), wUF distances shows that; 1- *in vivo* depletion of the 13 highly competitive strains explains significantly more variance than *in vivo* depletion of the 13 highly sensitive strains in the matrix, 2- the variance explained by *in vivo* depletion of community members' remains low and marginally significant in the roots (**Figure 38**, **panel-B**). The analysis beta-diversity indicates that perturbation by community members' depletions have a very limited effect on the structure bacterial communities under tested conditions. Taken together, the *A. thaliana* root-associated bacterial microbiota shows resilience against community members' depletions. These data point to the hypothesis that either the host leverages the establishment of the associated microbiota or the interactions between community members. It is also not excluded that identified communities members as highly competitive or sensitive have less important role *in planta*.

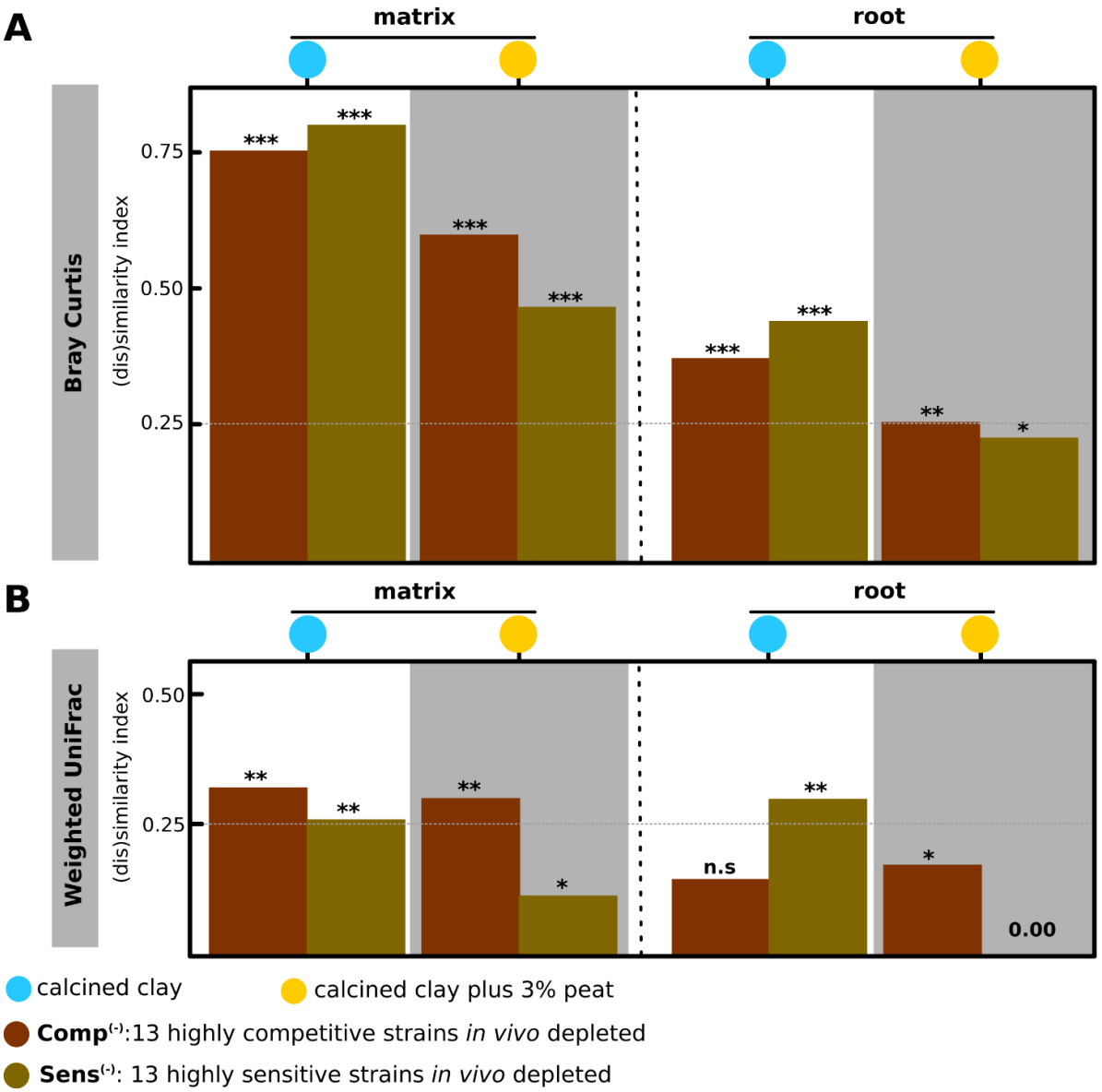


Figure 37| **Community members' depletion has a minor effect on the structure of roots bacterial communities.**

Panel-A and **-B** depict the analysis of similarity between perturbed communities and full communities computed on Bray Curtis (BC) and weighted UniFrac (wUF) distance metrics, respectively. The analysis of similarity tests for significant differences between the structure of Comp$^{(-)}$ communities and full communities *in silico* depleted from highly competitive bacteria and between Sens$^{(-)}$ communities and full communities *in silico* depleted from highly sensitive bacteria. Threshold for significance in (dis)similarity index is >0.25 for *p-value* <0.05. **Panel-A**, the analysis of similarities based on BC distance metric shows that both *in vivo* community members' depletions have a weaker effect on the community structure in the roots than the matrix. **Panel-B**, the analysis of similarities based on wUF distance metric indicates that *in vivo* depletion of the 13 highly competitive or sensitive have also a minor effect on the roots bacterial communities.

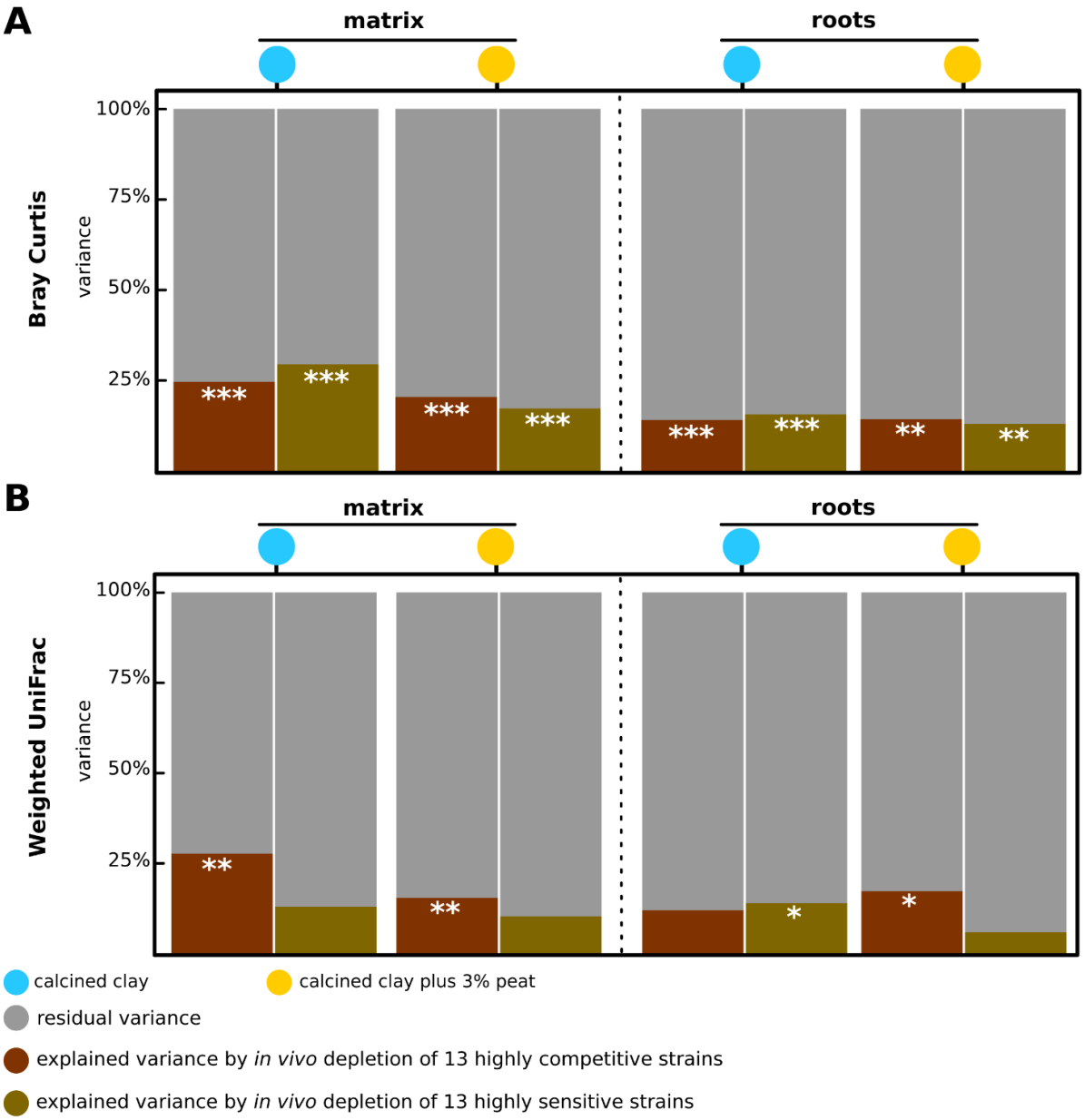


Figure 38| ***In vivo* depletion of community members' alters weakly the community structure *in planta*.**

Panel-A and **-B** show permutational multivariate analysis of variance of Bray Curtis (BC) and weighted UniFrac (wUF) distances between perturbed communities and full communities, respectively. Color code in the stack bars indicates residual variance, variance explained by *in vivo* depletion of the 13 highly competitive strians or variance explained by *in vivo* depletion of the 13 highly sensitive strians. 999 permutations were used to determine significance level of explained variance. **Panel-A**, the analysis of variance based on BC distance indicates that *in vivo* depletion of the 13 highly competitive strains explains less variance in roots than in matrix. **Panel-B**, the analysis of wUF distance indicates that both depletions have a marginal effect on the structure of root-associated bacteria.

Perturbations by community members' depletion alters the relative abundance of several bacterial isolates. *In vivo* depletion of the 13 highly competitive or sensitive strains had no effect on the bacterial community diversity and a very limited effect on the community structure. These observations indicate that the relative abundance of the majority of isolates, but unlikely all, is unchallenged upon the perturbations. Therefore, it is plausible that few isolates significantly increase or decrease in the relative abundance upon community members' depletions. To identify the isolates that are significantly increase or decrease in abundance upon the perturbations, I performed an enrichment. I compared $\log_2$ fold

change in the relative abundance of the isolates between perturbed communities (*in vivo* depletion of the 13 highly competitive or sensitive strains) in the matrix or the roots *versus* unperturbed communities (*in silico* depletion of highly competitive or sensitive strains, respectively) in the corresponding habitat (**Figure 39**). The bacterial isolates depicted in the heatmap correspond to the strains that show a significant increase or decrease in the relative abundance in at least one of the conditions. Although the community structure is marginally altered, several bacterial isolates are significantly show a significant increase or decrease in abundance upon community members' depletions (**Figure 39, panel-A** and **panel-B**). More precisely, *in vivo* depletion of the 13 highly sensitive strains leads to a significant increase in the relative abundance of 5 strains and a significant decrease in the relative abundance of 20 isolates. These results are consistent with the liquid microcosms and suggest that depletion of the highly sensitive community members' mainly restrict the growth of numerous isolates (**Figure 39, panel-A**). In contrast, *in vivo* depletion of the 13 highly competitive strains leads to a significant increase in the relative abundance of 19 isolates and a significant decrease in the relative abundance of 18 isolates (**Figure 39, panel-B**). Among the bacterial isolates that show a significant decrease in the relative abundance, 7 strains (n°329, 179, 170, 83, 695, 554 and 810) are shared between both *in vivo* depletions, these isolates are mainly *Proteobacteria*. Interestingly, only one isolate shared between both *in vivo* depletions (n°219 an *Acidovorax* strains) significantly increase in the relative abundance. In accordance with liquid microcosm findings, more bacterial isolates show a significant increase in the relative abundance upon *in vivo* depletion of the 13 highly competitive strains than upon *in vivo* depletion of the 13 highly sensitive strains. This is particularly clear in the calcined clay system supplemented with 3% peat. These data suggest that the highly competitive community members inhibit the growth of numerous and phylogenetically unrelated bacterial strains preferentially under carbon-rich conditions. The removal of competitive community members may free micro-habitat that be colonized by other community members. Consistently, several *Actinobacteria* strains and including four abundant soil bacteria show a significant increase in the relative abundance (*vs. in silico* depleted full community) in the roots but not in the surrounding calcined clay amended with 3% peat (**Figure 39, panel-B**). It is therefore plausible that the highly competitive strains out-compete abundant soil bacteria during root colonization. This result is reminiscent of the fact that *Actinobacteria*, and particularly abundant soil bacteria, are out-competed by other microbiota members in our ABBA screen (**Figure 15**). Among these bacterial isolates that show a significant increase in the relative abundance upon *in vivo* depletion of competitive strains, four of these isolates (n°101, 179, 772, 219) also have shown a significant increase in the relative abundance in at least one of the tested liquid microcosms (**Figure 31**). Altogether, both *in vivo* depletions are reported to alter the relative abundance of several bacterial strains, but the relative abundance of many more isolates significantly increase upon *in vivo* depletion of the 13 highly competitive strains than upon *in vivo* depletion of the 13 highly sensitive strains. These data corroborate the findings obtained for liquid microcosms and indicate that the highly competitive community members restrict the growth of other bacteria.

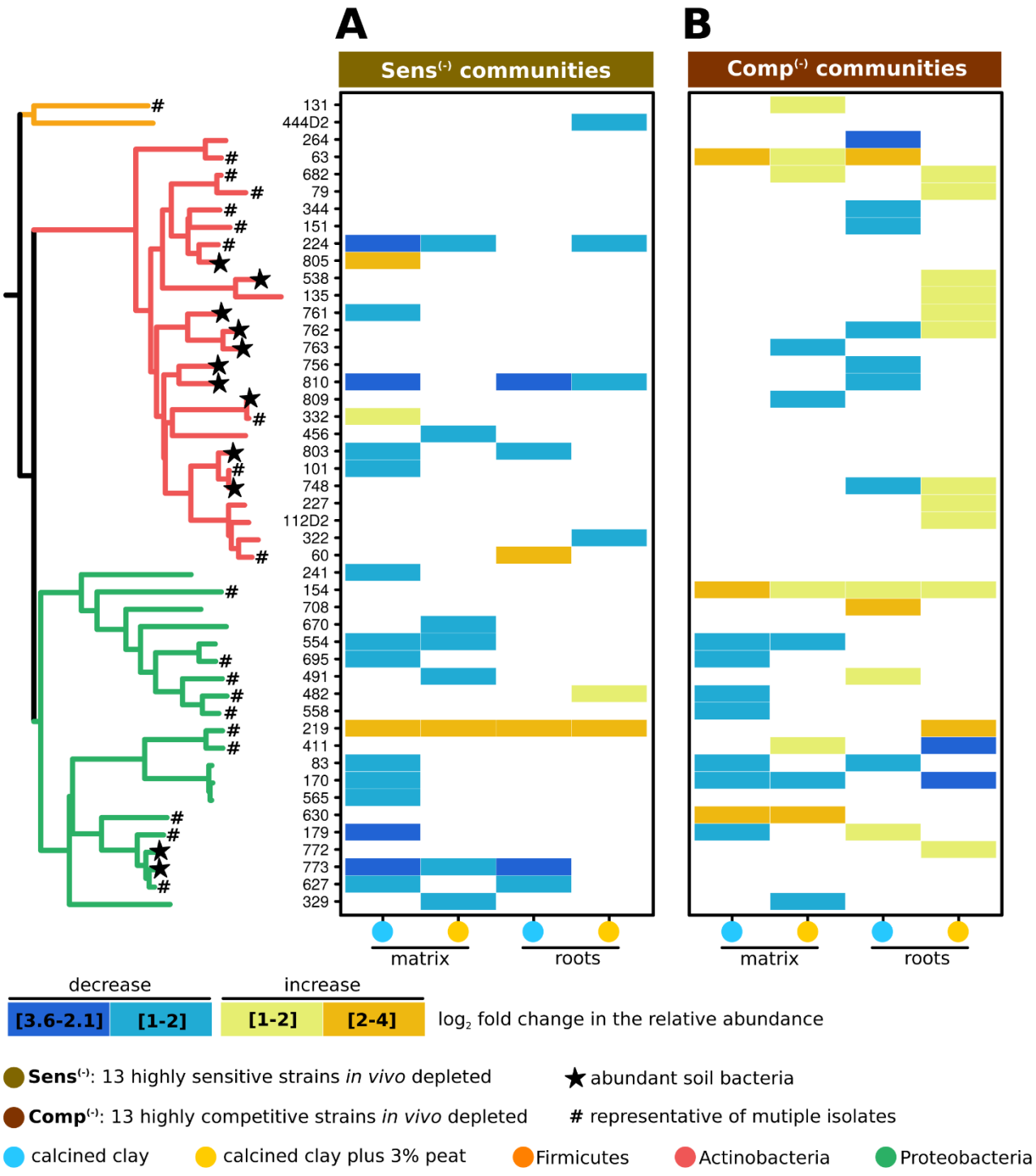


Figure 39| **Community members' depletions alter the abundance of bacterial isolates.**

The heatmap depicts log_2 fold change in the relative abudance of the bacterial isolates that significantly increase or decrease in abundance upon perturbations. Bacterial isolates are indicated in y-axis. Star indicates abundant soil bacteria and hash indicates representative isolates. The bacterial isolates in y-axis are ordered according to their phylogeny. Color code in the tree indicates phylum. Significant decrease in the relative abundance is color coded in a blue gradient and significant increase in the relative abundance is color coded in a red gradient. **Panel-A**, *in vivo* depletion of the 13 highly sensitive bacteria leads mainly to a significant decrease in the relative abundance of bacterial isolates. More bacterial isolates decrease in relative abundance in the matrix of calcined clay upon the depletion. **Panel-B**, *in vivo* depletion of the 13 highly competitive strains leads to a significant increase in the relative abundance of several *Actinobacteria* isolates in the roots in calcined clay plus 3% peat.

III.D. Discussion

Bacteria are important members of the plant root microbiota and have a critical role in plant growth and health (Hacquard *et al.*, 2015). Indeed, while plant growth-promoting bacteria help the host to elevate abiotic stresses, biocontrol bacterial strains contribute to the protection of the host against invading pathogens (Müller *et al.*, 2016, Whipps, 2001). Advances in sequencing technologies has led to an explosion of descriptive studies that comprehensively described the factors (soil type, host species, season, biogeography... etc.) that shape the bacterial root microbiota. While our knowledge on the composition and the function of the root-associated bacterial communities is extending faster than ever, our understanding of the fundamental principles that govern their assembly and stability remains fragmented. Importantly, high-throughput microbial profiling was critical in revealing that root-associated bacterial communities almost entirely derive from the surrounding soil biome. The establishment of the root-associated bacteria is initiated by the alteration of the soil at the vicinity of the roots (rhizosphere) via root exudation (Bulgarelli *et al.*, 2013). Interestingly, the examination of the *A. thaliana* root microbiota grown in different soils has clearly indicated that plants assemble a reproducible microbiota at the phylum level, largely dominated by three phyla *Proteobacteria*, *Actinobacteria* and *Bacteroidetes* (Bulgarelli *et al.*, 2013, Lundberg *et al.*, 2012). However at the genus level, the bacterial community composition remains highly dependent on the soil type, the plant species, or the residence time (Schlaeppi *et al.*, 2013, Dombrowski *et al.*, 2016). While it is well established that soil chemical and physical properties determine resident soil microbiota and that the plant-derived organic molecules initiate microbiota differentiation (Bulgarelli *et al.*, 2013), the fundamental role of bacteria-bacteria interactions, and more precisely of competitive interactions, in the establishment of the root microbiota remains unknown. This is mainly due to; 1- the complexity of soil and root-associated bacterial communities, 2- the multiple and confounding factors driving community establishment, and 3- the environmental noise that renders studying the role of inter-bacterial interactions in a community context a nontrivial task. A strategy to overcome these challenges is to reconstitute the bacterial microbiota of the plant roots under strictly controlled laboratory conditions. The establishment of simpler, yet highly representative, synthetic bacterial communities is a prerequisite to tackle the role of inter-bacterial interactions in shaping community diversity and stability (Bai *et al.*, 2015, Lebeis *et al.*, 2015). Recently, it has been shown in our research group that up to 66% of the bacterial OTUs detected in *A. thaliana* roots using a culture independent approach have a representative isolate in the culture collection. This finding indicates that a large majority of *A. thaliana* root-associated bacteria can be cultured (Bai et al. 2015). By combining culture collections that largely resemble the natural root microbiota with gnotobiotic systems and germ-free plants, it becomes conceivable to subject ecological hypotheses to experimental testing. Several studies have successfully employed reduced in complexity bacterial community, yet representative of the plant microbiota, to show the role of plant genotype or immune system in the establishment of host-associated bacterial communities or reveal that bacteria are better adapted to their cognate plant organ (Bai *et al.*, 2015, Lebeis *et al.*, 2015, Bodenhausen *et al.*,

2014, Niu *et al.*, 2017).

In the current chapter, I have explored the role of bacteria-bacteria interactions in altering the community composition and structure by using *A. thaliana* root-derived and abundant soil bacteria (Bai *et al.*, 2015). To this end, I have perturbed the bacterial communities by omitting the inoculation of either 13 highly competitive or 13 highly sensitive strains. The bacterial candidates for both depletions have been chosen based on their competitiveness or sensitiveness according to the study described in the previous chapter. In order to quantify the effect of the perturbations on the bacterial species richness and the community structure, I performed *in vivo* depletion of former or latter community members' from a full community that encompass 198 isolates. Perturbation by community members' depletion experiments were performed in liquid microcosms and in gnotobiotic systems. Based on the results obtained from the liquid microcosms, *in vivo* depletion of the 13 highly competitive strains had strongly altered the community diversity and structure, whereas *in vivo* depletion of 13 highly sensitive bacteria had only a limited effect. In contrast to liquid microcosms, both *in vivo* community members' depletions had no effect on the community diversity *in planta* or in the matrix and the structure of root-associated bacterial communities was weakly altered upon both perturbations. These data indicate that the highly competitive strains have a critical function for the assemblage of bacterial communities under aquaponic conditions and a more subtle/local effect *in planta* or in the soil-like matrix. However, a significant effect have been reported upon depletion of the highly competitive strains *in planta* when the clay matrix was amended with 3% peat. The result suggest that competitive strains defined based on the ABBA screen may require a carbon-rich habitat to increase their competitiveness. This hypothesis could be explained by the fact that the production of antimicrobials can have energetic cost for the cell that could be solved in carbon-rich habitat. The significant increase in the relative abundance of other community members' with a lower competitive potential (*i.e. Actinobacteria* isolates) in the roots suggests that the highly competitive strains may produce antimicrobials in the vicinity of roots that could be an additional factor of the differentiation of root-associated bacterial communities from the soil biome. These findings are indicative that competitive interactions between bacteria contribute to the establishment of the root-associated bacteria.

In vitro depletion experiments included different microcosms that analogously recapitulate limiting-growth condition and rich in nutrients growth condition. Remarkably, the *in vivo* depletion of the 13 highly competitive strains had a stronger effect on the bacterial community composition and structure than the *in vivo* depletion of 13 highly sensitive strains. Indeed, *in vivo* depletion of the former bacterial strains led to a sharp drop in species richness and a drastic shift in the community structure in almost all tested liquid microcosms. Moreover, depletion of the highly competitive members led to few *Proteobacteria* isolates, and more particularly the *Pseudomonas* strain n°9, to over-dominate the community (4 to 8 $\log_2$ fold increase in relative abundance). These data indicate that the highly competitive strains have a critical role in maintaining the community diversity in liquid microcosms under tested conditions. Although counter-intuitive, this phenomenon is consistent with theoretical-

based studies that have reported that competitive interactions promote community diversity and stability (Coyte *et al.,* 2015, Vetsigian *et al.,* 2011, Czárán *et al.* 2001). To the best of my knowledge, this study is unprecedented since the employed bacterial strains are phylogenetically diverse and reproduce a simpler, yet highly representative of the *A. thaliana* root microbiota. Nonetheless, it is important to note that although diverse bacterial isolates have been used, *Proteobacteria* and *Actinobacteria* were over-represented compared to *Firmicutes* and more importantly to *Bacteroidetes*. Therefore, it important to extend the present study to more isolates from both latter phyla in the future. It is also important to highlight that the highly competitive strains were mainly *Proteobacteria* isolates and the highly sensitive members were mainly *Actinobacteria* isolates, although former and latter isolates were chosen based on their competitiveness, it remains to be further tested whether *in vivo* depletion of highly competitive or highly sensitive members belonging to a same bacterial phylum leads to similar herein presented conclusions. The above-proposed experiments will shed light on the role of inter- and intra-phylum competitive interactions in promoting community diversity and stability. Importantly, it is important in future studies to follow the dynamics of bacterial communities over time since it is plausible that the community could shift from stability/instability over time.

A key finding in the presented study highlights the role of highly competitive members in promoting community diversity in liquid microcosms. Interestingly, the present is corroborated by computationally-inferred modeling studies that have predicted competition as a promoter of community diversity and stability (Coyte *et al.,* 2015, Vetsigian *et al.,* 2011, Czárán *et al.* 2001). Although increased diversity can jeopardy a community stability (Becker *et al.,* 2012), competitive interactions induce spatial segregation between competing species that result in coexisting species that stabilize community (kerr *et al.,* 2002, Kim *et al.,* 2008, Kelsic *et al.,* 2015). Moreover, competition in multi-species communities can be moderated by physical interposition of resistant members between a producer and a sensitive (Zapién-Campos et al., 2015McNally *et al.,* 2016, Gerardin *et al.,* 2016) or by chemical interposition through the modulation of produced antimicrobials within the community (Abrudan *et al.,* 2015). An additional speculative explanation, highly competitive bacteria are *bona fide* community members that dampen community members' that tend to over-dominate the community and/or negatively alter its diversity. However, it remains to be tested whether the exerted control on the community members is a direct effect through interference competition or an indirect effect through exploitative interactions. Although it is not yet clear by which mechanisms highly competitive bacteria maintain the community diversity, the current data are joined by the original observations reported by Niu *et al.,* 2017. The authors have explored the role of inter-bacterial interactions in community assemblage of a seven-species consortium. The seven species are a simplified community of the maize root microbiota. The authors tested several synthetic communities where each community was lacking a different member than the previous. By combining all possibilities, they inoculated these communities in axenic system and monitored the maize root-recolonization by selective culture-dependent approach. Interestingly, the removal of one particular member, *Enterobacter cloacae*, altered drastically the

community structure and led to the dominance of *Curtobacterium pusillum.* The authors' and the present study point to the assumption that bacterial communities have “keystone species” that preserve the community diversity and stability. Under the present study, “keystone” community members are the highly competitive strains that likely promote the diversity and stability of the community through competitive interactions. Interestingly, the concept of keystone have been also extended to multiple kingdoms communities as shown by Agler *et al.,* 2016. The authors have followed bacterial, fungal and oomycetal community dynamics after abiotic factors, host genotype and pathogen colonization have been manipulated under field and laboratory conditions. A key finding of the study was the identification of a subset of microbes that act as “hub” community members. Particularly, two “hub” microbes, an obligate biotrophic oomycete pathogen *Albugo* and a yeast *Dioszegia*, were further analyzed. Interestingly, identified “hub” microbes were strongly interconnected to community members and have been shown to alter leaf microbiota. Therefore, manipulating “hub” microbes will undoubtedly have consequences on microbial communities and subsequently on the associated host (Cottee-Jones and Whittaker, 2012, Trosvik and de Muinck, 2015).

Mounting evidence points toward that “keystone species” are important for the stability of microbial communities (Ze *et al.,* 2012, Fisher and Mehta, 2014, Trosvik and de Muinck, 2015, Niu *et al.,* 2017). However, our results suggest that the structure of the root-associated bacteria under gnotobiotic conditions is largely robust, despite the applied perturbations *via* community members' depletion. The present finding is in incongruity with Niu *et al.,* 2017 study and more remarkably with liquid microcosms. Although, the isolates selected in this study do not constitute the entire *Arabidopsis thaliana* roots microbiota, it is important to highlight that this study included 198 taxonomically diverse bacterial strains, far more complex community than the seven-member community used by Niu *et al.,* 2017. It is then plausible that a “keystone species” in a low complexity community may not be a “keystone species” in a highly complex community where functional redundancy between the community members is expected. As for the discrepancy between liquid microcosms and gnotobiotic growth system, the former experimental corresponds to aquaponic growth cultures where strong and multiple interactions between bacteria are more likely to play an important role in altering the community diversity and structure. Indeed, two bacterial cells are more likely to be affected by antimicrobials in liquid culture than in a soil-like matrix or on the rhizoplane. Although physically separated, bacteria-bacteria interactions are facilitated through the diffusion of secreted molecules in liquid microcosm. Therefore, all community members can virtually interact with all other members in liquid microcosm (higher inter-bacteria connectivity). In contrast, inter-bacterial interactions in gnotobiotic systems are constrained by the matrix moisture and the proximity of interacting partners. Secreted molecules may face obstacles to diffuse through clay pores in order to reach target species. It is then plausible that interactions between bacteria in the matrix or the roots have a stronger impact “locally” at the micro-scale that cannot be detected by profiling the entire community. The fact that several bacterial isolates increase in the relative abundance upon depletion of the highly competitive

bacteria supports this hypothesis in the clay system amended with 3% peat. More importantly, highly competitive and sensitive bacteria have been defined based on a carbon-rich synthetic growth medium. Medium composition is an important factor for the secretion of antimicrobial metabolites. It is not excluded that other community members become highly competitive in the roots or that the production of antimicrobials is slightly different from those produced by bacteria in the synthetic medium. On the other hand, the plant is a living experimental system and host-microbe interactions are as plausible as microbe-microbe interactions. Therefore, it is also not excluded that the plant leverages the establishment of the root-associated bacterial microbiota and contribute to the maintenance of the community homeostasis. However, more empirical evidences are required to uncouple the role of bacteria-bacteria interactions from bacteria-host interaction in the establishment of the root microbiota. In future investigations, it is important to include time series resolution in the profiling of microbial communities. Following the dynamics of the microbiota is important to ascertain that established communities are stable/unstable over time. Also, since the depletion of the 13 highly competitive strains had strongly altered the community structure liquid microcosms but only weakly *in planta*, it remains to be tested whether the depletion of more bacterial isolates will significantly alter the root-associated bacterial communities. These additional experiments are needed to identify the minimal number of depletion of both highly competitive and highly sensitive strains that lead to the community collapse *in planta*. Least but not last, it might be relevant to perform random members' depletion in order to overcomes taxonomic bias and evaluate whether the phylogeny of the depleted bacteria play a more consequent role on the community composition than their actual competitiveness.

To resume, the data presented in this chapter indicate that bacteria-bacteria interactions are important for the establishment of bacterial communities. The study of perturbed communities in liquid microcosms have showed that the highly sensitive strains and the highly competitive strains have different roles in the community. The former group of strains may rather promote the growth of several other community members suggesting that cooperative interactions are alternative mechanisms to persist in a multi-species community. In contrast, the latter group of strains promoted the community diversity through likely the secretion of antimicrobials that restrict the over-growth of several other community members, which highlights the potential self-organizational properties of the bacterial microbiota. Competitive inter-bacterial interactions have more subtle effect *in planta* and in the clay matrix indicating that these interactions are likely more important at the micro-scale. Nonetheless, *in planta* perturbation experiments suggest that bacterial competitiveness may be more important in carbon-rich habitats, since the depletion of the highly competitive strains led to the enrichment of several *Actinobacteria* strains in the roots. The present study provides indicative results for the eminence role of inter-bacterial interactions through the secretion of antimicrobials in promoting the microbial communities diversity and stability in nature.

III.E. References

- Abrudan, Monica I., Fokko Smakman, Ard Jan Grimbergen, Sanne Westhoff, Eric L. Miller, Gilles P. van Wezel, and Daniel E. Rozen. "Socially Mediated Induction and Suppression of Antibiosis during Bacterial Coexistence." *Proceedings of the National Academy of Sciences* 112, no. 35 (September 1, 2015): 11054–59. doi:10.1073/pnas.1504076112.
- Agler, Matthew T., Jonas Ruhe, Samuel Kroll, Constanze Morhenn, Sang-Tae Kim, Detlef Weigel, and Eric M. Kemen. "Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation." *PLOS Biology* 14, no. 1 (January 20, 2016): e1002352. doi:10.1371/journal.pbio.1002352.
- Anderson, Marti J. "A New Method for Non-Parametric Multivariate Analysis of Variance." *Austral Ecology* 26, no. 1 (February 1, 2001): 32–46. doi:10.1111/j.1442-9993.2001.01070.pp.x.
- Anderson, Marti J., and Daniel C. I. Walsh. "PERMANOVA, ANOSIM, and the Mantel Test in the Face of Heterogeneous Dispersions: What Null Hypothesis Are You Testing?" *Ecological Monographs* 83, no. 4 (November 1, 2013): 557–74. doi:10.1890/12-2010.1.
- Bai, Yang, Daniel B. Müller, Girish Srinivas, Ruben Garrido-Oter, Eva Potthoff, Matthias Rott, Nina Dombrowski, et al. "Functional Overlap of the Arabidopsis Leaf and Root Microbiota." *Nature* 528, no. 7582 (December 17, 2015): 364–69. doi:10.1038/nature16192.
- Becker, Joachim, Nico Eisenhauer, Stefan Scheu, and Alexandre Jousset. "Increasing Antagonistic Interactions Cause Bacterial Communities to Collapse at High Diversity." *Ecology Letters* 15, no. 5 (May 2012): 468–74. doi:10.1111/j.1461-0248.2012.01759.x.
- Bodenhausen, Natacha, Miriam Bortfeld-Miller, Martin Ackermann, and Julia A. Vorholt. "A Synthetic Community Approach Reveals Plant Genotypes Affecting the Phyllosphere Microbiota." *PLOS Genetics* 10, no. 4 (avr 2014): e1004283. doi:10.1371/journal.pgen.1004283.
- Bulgarelli, Davide, Matthias Rott, Klaus Schlaeppi, Emiel Ver Loren van Themaat, Nahal Ahmadinejad, Federica Assenza, Philipp Rauf, et al. "Revealing Structure and Assembly Cues for Arabidopsis Root-Inhabiting Bacterial Microbiota." *Nature* 488, no. 7409 (August 2, 2012):91–95. doi:10.1038/nature11336.
- Bulgarelli, Davide, Klaus Schlaeppi, Stijn Spaepen, Emiel Ver Loren van Themaat, and Paul Schulze-Lefert. "Structure and Functions of the Bacterial Microbiota of Plants." *Annual Review of Plant Biology* 64, no. 1 (2013): 807–38. doi:10.1146/annurev-arplant-050312-120106.
- Cottee-Jones, Henry Eden W., and Robert J. Whittaker. "Perspective: The Keystone Species Concept: A Critical Appraisal." *Frontiers of Biogeography* 4, no. 3 (January 1, 2012). http://escholarship.org/uc/item/15d2b65t.

- Coyte, Katharine Z., Jonas Schluter, and Kevin R. Foster. "The Ecology of the Microbiome: Networks, Competition, and Stability." *Science* 350, no. 6261 (November 6, 2015): 663–66. doi:10.1126/science.aad2602.
- Czárán, Tamás L., Rolf F. Hoekstra, and Ludo Pagie. "Chemical Warfare between Microbes Promotes Biodiversity." *Proceedings of the National Academy of Sciences* 99, no. 2 (January 22, 2002):786–90. doi:10.1073/pnas.012399899.
- Dombrowski, Nina, Klaus Schlaeppi, Matthew T. Agler, Stéphane Hacquard, Eric Kemen, Ruben Garrido-Oter, Jörg Wunder, George Coupland, and Paul Schulze-Lefert. "Root Microbiota Dynamics of Perennial Arabis Alpina Are Dependent on Soil Residence Time but Independent of Flowering Time." *The ISME Journal* 11, no. 1 (January 2017): 43–55. doi:10.1038/ismej.2016.109.
- Edwards, Joseph, Cameron Johnson, Christian Santos-Medellín, Eugene Lurie, Natraj Kumar Podishetty, Srijak Bhatnagar, Jonathan A. Eisen, and Venkatesan Sundaresan. "Structure, Variation, and Assembly of the Root-Associated Microbiomes of Rice." *Proceedings of the National Academy of Sciences* 112, no. 8 (February 24, 2015): E911–20. doi:10.1073/pnas.1414592112.
- Fierer, Noah, Mark A. Bradford, and Robert B. Jackson. "Toward an Ecological Classification of Soil Bacteria." *Ecology* 88, no. 6 (June 2007): 1354–64.
- Fisher, Charles K., and Pankaj Mehta. "Identifying Keystone Species in the Human Gut Microbiome from Metagenomic Timeseries Using Sparse Linear Regression." *PLOS ONE* 9, no. 7 (juil 2014): e102451. doi:10.1371/journal.pone.0102451.
- Friesen, Maren L., Stephanie S. Porter, Scott C. Stark, Eric J. von Wettberg, Joel L. Sachs, and Esperanza Martinez-Romero. "Microbially Mediated Plant Functional Traits." *Annual Review of Ecology, Evolution, and Systematics* 42, no. 1 (2011): 23–46. doi:10.1146/annurev-ecolsys-102710-145039.
- Gerardin, Ylaine, Michael Springer, and Roy Kishony. "A Competitive Trade-off Limits the Selective Advantage of Increased Antibiotic Production." *Nature Microbiology* 1 (September 26, 2016):16175. doi:10.1038/nmicrobiol.2016.175.
- Hacquard, Stéphane, Ruben Garrido-Oter, Antonio González, Stijn Spaepen, Gail Ackermann, Sarah Lebeis, Alice C. McHardy, et al. "Microbiota and Host Nutrition across Plant and Animal Kingdoms." *Cell Host & Microbe* 17, no. 5 (May 13, 2015): 603–16. doi:10.1016/j.chom.2015.04.009.
- Herrera Paredes, Sur, and Sarah L. Lebeis. "Giving back to the Community: Microbial Mechanisms of Plant–soil Interactions." *Functional Ecology* 30, no. 7 (July 1, 2016): 1043–52. doi:10.1111/1365-2435.12684.

- Kelsic, Eric D., Jeffrey Zhao, Kalin Vetsigian, and Roy Kishony. "Counteraction of Antibiotic Production and Degradation Stabilizes Microbial Communities." *Nature* 521, no. 7553 (May 28, 2015): 516–19. doi:10.1038/nature14485.
- Kerr, Benjamin, Margaret A. Riley, Marcus W. Feldman, and Brendan J. M. Bohannan. "Local Dispersal Promotes Biodiversity in a Real-Life Game of Rock–paper–scissors." *Nature* 418, no.6894 (July 11, 2002): 171–74. doi:10.1038/nature00823.
- Kim, Hyun Jung, James Q. Boedicker, Jang Wook Choi, and Rustem F. Ismagilov. "Defined Spatial Structure Stabilizes a Synthetic Multispecies Bacterial Community." *Proceedings of the National Academy of Sciences* 105, no. 47 (November 25, 2008): 18188–93. doi:10.1073/pnas.0807935105.
- Lebeis, Sarah L., Sur Herrera Paredes, Derek S. Lundberg, Natalie Breakfield, Jase Gehring, Meredith McDonald, Stephanie Malfatti, et al. "PLANT MICROBIOME. Salicylic Acid Modulates Colonization of the Root Microbiome by Specific Bacterial Taxa." *Science (New York, N.Y.)* 349, no. 6250 (August 21, 2015): 860–64. doi:10.1126/science.aaa8764.
- Lozupone, Catherine, Manuel E Lladser, Dan Knights, Jesse Stombaugh, and Rob Knight. "UniFrac: An Effective Distance Metric for Microbial Community Comparison." *The ISME Journal* 5, no. 2 (February 2011): 169–72. doi:10.1038/ismej.2010.133.
- Lundberg, Derek S., Sarah L. Lebeis, Sur Herrera Paredes, Scott Yourstone, Jase Gehring, Stephanie Malfatti, Julien Tremblay, et al. "Defining the Core Arabidopsis Thaliana Root Microbiome." *Nature* 488, no. 7409 (August 2, 2012): 86–90. doi:10.1038/nature11237.
- McNally, Luke, Eryn Bernardy, Jacob Thomas, Arben Kalziqi, Jennifer Pentz, Sam P. Brown, Brian K. Hammer, Peter J. Yunker, and William C. Ratcliff. "Killing by Type VI Secretion Drives Genetic Phase Separation and Correlates with Increased Cooperation." *Nature Communications* 8 (February 6, 2017). doi:10.1038/ncomms14371.
- Müller, Daniel B., Christine Vogel, Yang Bai, and Julia A. Vorholt. "The Plant Microbiota: Systems-Level Insights and Perspectives." *Annual Review of Genetics* 50, no. 1 (2016): 211–34. doi:10.1146/annurev-genet-120215-034952.
- Niu, Ben, Joseph Nathaniel Paulson, Xiaoqi Zheng, and Roberto Kolter. "Simplified and Representative Bacterial Community of Maize Roots." *Proceedings of the National Academy of Sciences of the United States of America* 114, no. 12 (March 21, 2017): E2450–59. doi:10.1073/pnas.1616148114.
- Schlaeppi, Klaus, Nina Dombrowski, Ruben Garrido Oter, Emiel Ver Loren van Themaat, and Paul Schulze-Lefert. "Quantitative Divergence of the Bacterial Root Microbiota in Arabidopsis Thaliana Relatives." *Proceedings of the National Academy of Sciences* 111, no. 2 (January 14, 2014): 585–92. doi:10.1073/pnas.1321597111.

- Trosvik, Pål, and Eric Jacques de Muinck. "Ecology of Bacteria in the Human Gastrointestinal Tract-identification of Keystone and Foundation Taxa." *Microbiome* 3 (2015): 44. doi:10.1186/s40168-015-0107-4.
- Vetsigian, Kalin, Rishi Jajoo, and Roy Kishony. "Structure and Evolution of Streptomyces Interaction Networks in Soil and In Silico." *PLOS Biology* 9, no. 10 (October 25, 2011): e1001184. doi:10.1371/journal.pbio.1001184.
- Vorholt, Julia A. "Microbial Life in the Phyllosphere." *Nature Reviews Microbiology* 10, no. 12 (December 2012): 828–40. doi:10.1038/nrmicro2910.
- Wagner, Maggie R., Derek S Lundberg, Tijana G. del Rio, Susannah G. Tringe, Jeffery L. Dangl, and Thomas Mitchell-Olds. "Host Genotype and Age Shape the Leaf and Root Microbiomes of a Wild Perennial Plant." *Nature Communications* 7 (July 12, 2016). doi:10.1038/ncomms12151.
- Warton, David I., Stephen T. Wright, and Yi Wang. "Distance-Based Multivariate Analyses Confound Location and Dispersion Effects." *Methods in Ecology and Evolution* 3, no. 1 (February 1, 2012): 89–101. doi:10.1111/j.2041-210X.2011.00127.x.
- Whipps, J. M. "Microbial Interactions and Biocontrol in the Rhizosphere." *Journal of Experimental Botany* 52, no. Spec Issue (March 2001): 487–511.
- Zapién-Campos, R., Olmedo-Álvarez, G., and Santillán, M. (2015). Antagonistic interactions are sufficient to explain self-assemblage of bacterial communities in a homogeneous environment: a computational modeling approach. *Front Microbiol* 6, 489. doi:10.3389/fmicb.2015.00489.
- Ze, Xiaolei, Sylvia H Duncan, Petra Louis, and Harry J Flint. "Ruminococcus Bromii Is a Keystone Species for the Degradation of Resistant Starch in the Human Colon." *The ISME Journal* 6, no.8 (August 2012): 1535–43. doi:10.1038/ismej.2012.4.

Concluding remarks

Competitive interactions mediated by antimicrobials are not only important for the host, but also for the host-associated microbial community, diversity (Czárán et al., 2002), spatial structure (Narisawa *et al.*, 2008) and stability (Coyte et al., 2015). To what extent antagonistic interactions contribute to the establishment of the plant-associated microbiota is poorly understood. In this study, I explored the competitive potential of several *A. thaliana* root-associated bacteria in order to define highly competitive community members and test their role in altering the community diversity and structure using different experimental systems. By combining genomic, metabolomic and phenotypic analyses, I explored the competitive potential of phylogenetic diverse bacterial strains. The analysis of the bacterial genomes revealed that several microbiota members harbor diverse biosynthetic gene clusters that encode enzymatic pathways for the production of antimicrobials, including bacteriocins, nonribosomal peptides, or polyketides. The root-associated bacteria harbor the genetic potential to engage in contact-independent competitive interactions. The analysis of the bacterial metabolites showed that several strains secrete genome-predicted antimicrobials. This studies joins others by indicating that most of BGCs are silent and not expressed under laboratory conditions (Rutledge and Challis, 2015). By screening for mutual inhibitions, I revealed that 66% of the isolates engage in antagonistic interactions, though these interactions represented 2.5% of overall tested interactions. The low frequency in antagonistic interactions indicated that resistance to antimicrobials is a widespread competitive mechanisms (Davies and Davies, 2010). Exploring the network of inhibitions allowed us to define two groups of strains with contrasting competitive potential; 13 highly competitive and 13 highly sensitive bacteria. To test the role of bacteria-bacteria interactions in altering the community diversity and structure of microbial communities, I perturbed phylogenetically-diverse synthetic bacterial community by *in vivo* depletion of the former or the latter groups of strains and used two different experimental systems; liquid microcosms and gnotobiotic *in planta* systems. The perturbation of bacterial communities in liquid microcosms revealed that highly competitive bacteria are important community members that promote the community diversity and stability. Interestingly, *in planta* root-associated microbiota showed a resilience against the applied perturbations. This study indicates that inter-bacterial competitive interactions are important for the community diversity and structure in niche-dependent manner and these interactions have rather a “local”, at the micro-scale, role *in planta.* More over, this study indicated that in a community highly competitive bacteria compete against phylogenetically diverse community memebers, whereas highly sensitive bacteria rather cooperate with community members.

References

- Coyte, K. Z., Schluter, J., and Foster, K. R. (2015). The ecology of the microbiome: Networks, competition, and stability. *Science* 350, 663–666. doi:10.1126/science.aad2602.
- Czárán, T. L., Hoekstra, R. F., and Pagie, L. (2002). Chemical warfare between microbes promotes biodiversity. *Proc Natl Acad Sci U S A* 99, 786–790. doi:10.1073/pnas.012399899.
- Davies, J., and Davies, D. (2010). Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.* 74, 417–433. doi:10.1128/MMBR.00016-10.
- Narisawa, N., Haruta, S., Arai, H., Ishii, M., and Igarashi, Y. (2008). Coexistence of antibiotic-producing and antibiotic-sensitive bacteria in biofilms is mediated by resistant bacteria. *Appl. Environ. Microbiol.* 74, 3887–3894. doi:10.1128/AEM.02497-07.
- Rutledge, P. J., and Challis, G. L. (2015). Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat. Rev. Microbiol.* 13, 509–523. doi:10.1038/nrmicro3496.

Materials and Methods

Bacterial culture collection. The culture collection employed comprises 198 isolates (**figure 40 Supplementary, Table 1**), 167 isolates are root-associated strains and 31 are abundant soil strains (Bai *et al.*, 2015). Root-associated bacteria were isolated from *Arabidopsis thaliana* roots that were grown in Cologne soil. Abundant soil bacteria were isolated from unplanted Cologne soil (Bai et al., 2016). These isolates cover four phyla "*Firmicutes*, *Actinobacteria*, *Proteobacteria* and *Bacteroidetes*" and 25 bacterial families. The genome of all isolates have been sequenced and downloaded from "www.at-sphere.com".

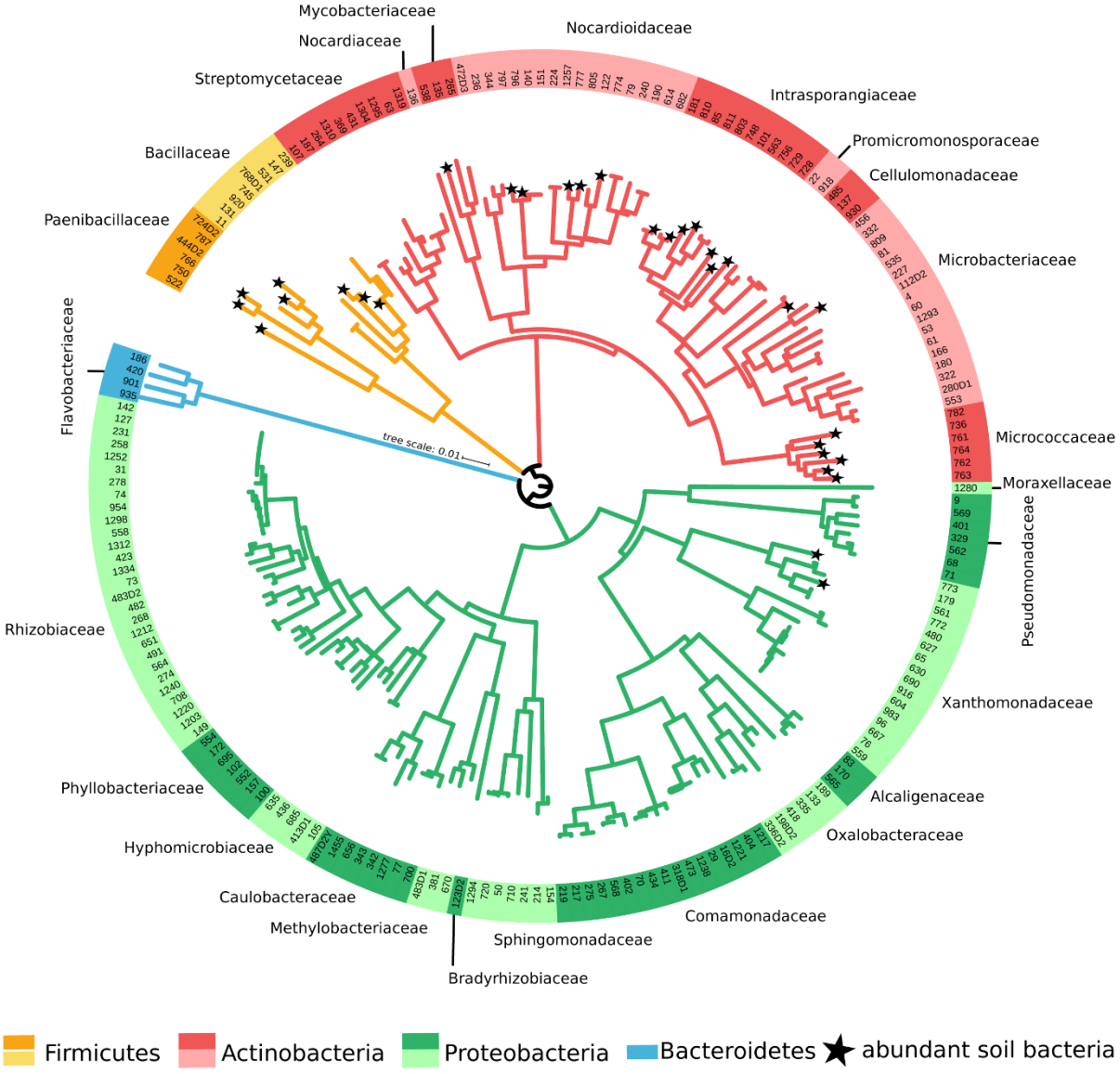


Figure 40| **Phylogenetically diverse bacterial strains used in this study.**

The circular phylogenetic tree shows 16s rRNA genes distances between 198 bacterial strains employed in this study (*i.e.* the analysis of genomes, metabolites and inter-bacterial antagonistic interactions and in community perturbations). The color code indicates the phylum and stars indicate abundant soil bacteria. The strains represent 25 bacterial families and four bacterial phyla. 167 strains are *Arabidopsis thaliana* root-derive bacteria and 31 strains are abundant soil bacteria.

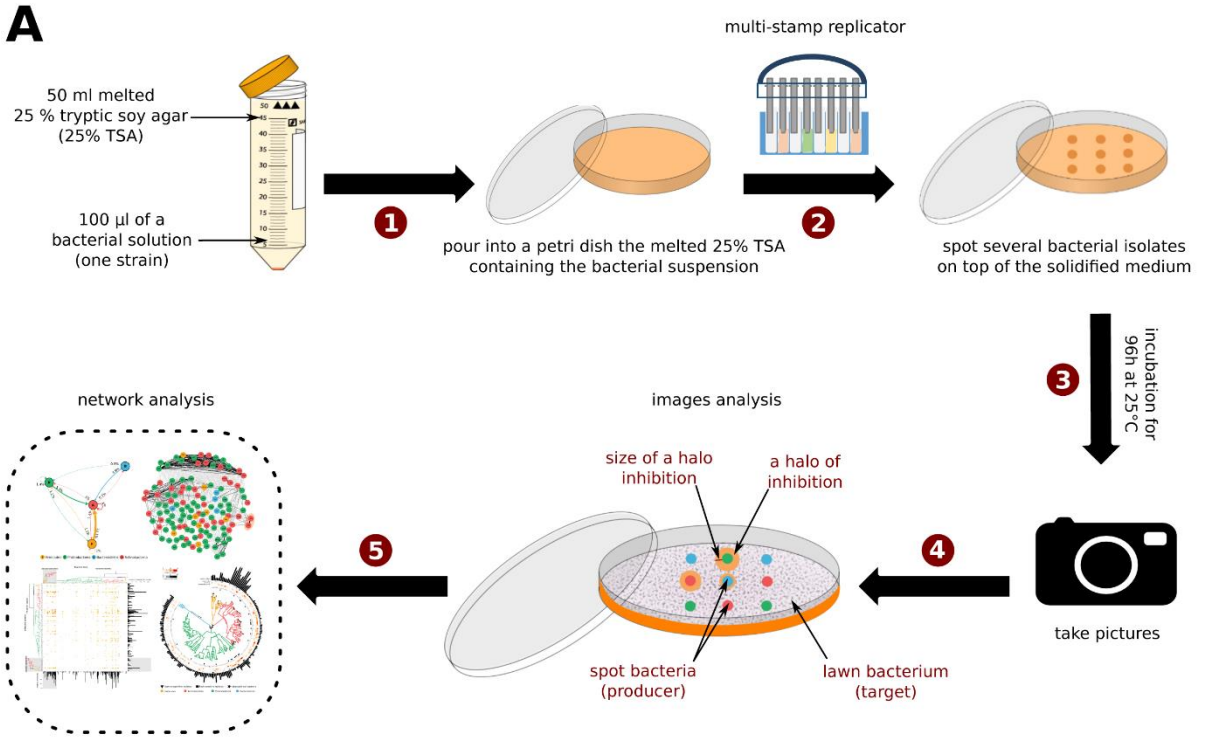
Genomes analysis for biosynthetic gene clusters. The bacterial genomes were submitted to "www.antismash.secondarymetabolites.org" (version 3.0, Madema *et al.*, 2011) in order to predict biosynthetic clusters. Output data from antiSMASH analysis were aggregated at BGC class for each isolate and are displayed in **Supplementary Table 2**. For the comparative BGCs analysis, data were aggregated at family or phylum level. For network analysis, a correlation coefficient was calculated between isolates under R statistical environment (R Development Core Team, 2008) using the R package "qgraph" (Epskamp *et al.*, 2012).

Metabolomic Analysis of 198 Bacterial Isolates. Each bacterial strain was grown separately in 25% tryptic soy agar (25% TSA) plate (25% of BBL™ Trypticase™ Soy, BD with 1.8% BactoAgar, BD). After 7 days of incubation at 25 °C, three to four agar plugs were taken from the periphery and inside of the bacterial colony. Agar plugs were crushed and washed with 500 µl of sterile water. Extraction followed subsequently in 500 µl ethyl acetate (EtOAc) and methanol (MeOH). Between each extraction step, samples were vortexed for 30-45 seconds. After each extraction, the solvents were evaporated and the residue was redissolved in 500 µl of MeOH, LC-MS grade and filtered through a 0.2 µm membrane into HPLC vials. Solvents for blanks (non-inoculated medium) were extracted according to above-described protocol. The analysis of metabolites was performed once. Samples were analyzed by HPLC-MS/MS on a micrOTOF-Q mass spectrometer (Bruker) with ESI-source coupled with a HPLC Dionex Ultimate 3000 (Thermo Scientific) using a Zorbax Eclipse Plus C18 1.8 µm column, 2.1x50 mm (Agilent). The column temperature was 45 degree Celsius. MS data were acquired over a range from 100-3000 m/z in positive mode. Auto MS/MS fragmentation was achieved with rising collision energy (35-50 keV over a gradient from 500-2000 m/z) with a frequency of 4 Hz for all ions over a threshold of 100. uHPLC begins with 90 % H_2O containing 0.1% acetic acid. The gradient starts after 0.5 min to 100% Acetonitrile (0.1% acetic acid) in 4 min. 2 ul of sample solution was injected to a flow of 0.8 ml/min. All MS/MS data were converted and transferred to the GNPS server (gnps.ucsd.edu) (Wang et al. Nat Biotechnol. 2016) and molecular networking was performed based on the GNPS data analysis workflow using the spectral clustering algorithm (Guthals *et al.*, 2012). Samples attributes were assigned to the data files (198 isolates, x families 4 phyla and two solvents). For the network analysis, all nodes that contained ions from blank medium were removed. The network was visualized via Cytoscape 3.3.0. For comparative analysis, data were aggregated at the phylum level. For principal coordinates analysis, data were exported through "Create Cluster Buckets" option on GNPS data analysis Advanced Output Options. Obtained table was used to calculate Sørensen index

using qiime bioinformatics pipeline (Caporaso *et al.*, 2013). Principal coordinates analysis was computed under R statistical environment (R Development Core Team, 2008) using the R package "phyloseq" (McMurdie and Holmes, 2013).

Screen for antagonistic inter-bacterial interactions. A descriptive flowchart of the screen is depicted in **Figure 41**. Bacterial isolates were cultured for 7 days in 25% tryptic soy broth (25% TSB). A bacterial solution of 100 µl was re-suspended in 50 ml melted 25% TSA and poured afterwards in a square petri dish (120x120 mm). After medium solidification, several bacterial isolates were spotted on top of the medium using multi-stamp replicator. Between each stamp, the replicator was sterilized by 70% EtOH followed by flaming. All bacterial handling was performed under sterile working conditions. Plates were indicated at 25°C. After 96h of incubation, pictures were taken and analyzed for halo of inhibition. The size of the halo of inhibition was measured using ImageJ (Schneider *et al.*, 2012). For comparative analysis, scores of inhibitions were aggregated at the phylum, family level or according to root-associated or abundant soil bacteria groups. Regarding network analysis of antagonistic activity or sensitivity profiles, a correlation coefficient was calculated between isolates under R statistical environment (R Development Core Team, 2008) using the R package "qgraph" (Epskamp *et al.*, 2012). The screen for 198 isolates was conducted once and validated by chosen a random set of isolates that have been re-screened for antagonistic activity as described above. The Screen for antagonistic activity against clinical isolates were performed at the Institute for Medicine, Microbiology, Immunology and Hygiene of Cologne. All bacterial isolates were screened once and bacterial isolates that showed antagonistic activity were re-screened twice more.

Bacterial growth and medium. Bacterial isolates were pre-cultured from frozen glycerol stock in 25% TSA for six to seven days at 25°C. Afterwards, a single colony, if possible, were picked and re-suspended in 25% TSB and incubated for additional six to seven days. These two steps of bacterial pre-growth were implemented as standard protocol prior *in vitro* and *in planta* perturbation experiments. Complex medium used in perturbation experiment corresponds to 25% tryptic soy broth. Minimal medium used in perturbation experiments corresponds to M9, minimal salts (Sigma), amended with trace elements, vitamin B solution and carbon source (glucose and fructose 1.64 g/l, saccharose 0.8 g/l, citric acid and lactic acid 0.64 g/l, succinic acid and serine 0.92g/l and glutamic acid and serine 0.8 g/l).



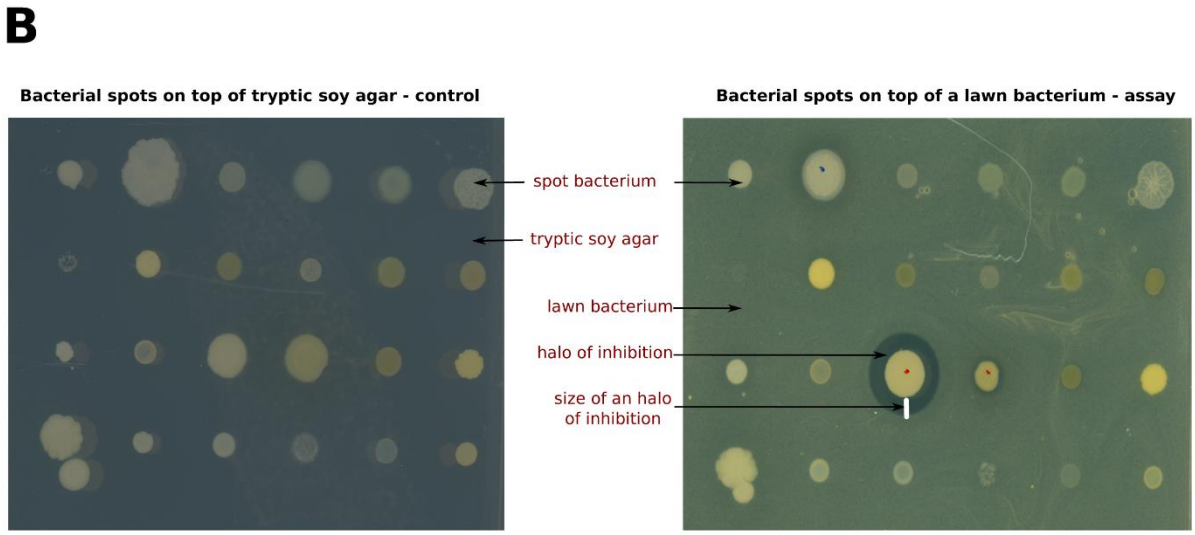


Figure 41| **Flowchart and output examples describing the antagonistic bacteria aacteria interactions assay.**

Panel-A, the flowchart describes the screen for inter-bacterial antagonistic interactions. The screen is indicated by the following steps; 1- 100 µl of a bacterial solution (one strain) is re-suspended in 50 ml of melted 25% tryptic soy agar (25% TSA) and poured into a petri dish. 2- several bacterial isolates are spotted on top of the solidified medium. 3- the plates are incubated during 96h at 25°C. 4- after the incubation period, pictures are taken and the size of halos of inhibitions are measured. 5-the data are collected for downstream network analysis. **Panel-B**, two pictures illustrating the screen output. Left-side picture shows spot of bacteria on top of 25% TSA. Right-side picture shows spot bacteria on top of a lawn bacterium. A halo of inhibition ca be detected from the right-side picture.

Community perturbation in liquid microcosm. Bacteria were pre-cultured seven days in 25% TSB, then pooled and washed twice with $MgCl_2$ 10 mM. Bacterial mix was resuspended in $MgCl_2$ 10 mM and OD_{600} was adjusted to 0.5 prior to the inoculation of 1 ml of bacterial solution into 50 ml of medium. For perturbation conditions, 13 highly competitive or sensitive isolates were omitted from the inoculum. For full community condition, both 13 highly

competitive and sensitive were inoculated. Bacterial communities were incubated at 25°C either under shaking condition (150 rpm) or standing condition in 250 ml glass flasks. After 96 h of incubation, communities were centrifuged at maximum speed during 10 minutes, resuspended in 1 ml Nuclease-Free water (Qiagen) and then snap-frozen in liquid nitrogen and stored at -80°C until DNA extraction.

DNA extraction from liquid microcosm. To extracted DNA from liquid microcosm, bacterial cultures were centrifuged for 10 minutes and resuspended in 1 ml Nuclease-Free water (Qiagen) and transferred to 2 ml tube Lysing Matrix E (MP Biomedicals) and snap-frozen in liquid nitrogen. Tubes were homogenized twice by Precellys 24 tissue lyser (Bertin Technologies) and the DNA extracted according to following. After homogenization, 180 µl of lysosyme was added to each tube and briefly vortexed prior incubation at 37°C for 30 minutes. After incubation time, 4 µl of Rnase and 20 µl of proteinase K were added to each tube. Tube were incubated at room temperature for 5 minutes and vortexed intermittently. Afterwards, DNA was extracted according to QIAamp DNA Micro Kit (Qiagen). DNA was eluted in Nuclease-Free water (Qiagen) and stored at -20°C until library preparation for amplicon sequencing.

Community perturbation in gnotobiotic system. Bacterial strains were cultured in 25% TSB in 96 well-plate for seven days, then pulled and washed twice with $MgCl_2$ 10 mM. Bacterial mix was resuspended in $MgCl_2$ 10 mM and OD_{600} adjusted to 0.5. For perturbation conditions, 13 highly competitive or sensitive isolates were omitted from the inoculum. For full community condition, both 13 highly competitive and sensitive were included in the bacterial solution. For inoculations, 1ml of bacterial solution with OD_{600} of 0.5 were resuspended into 70 ml half MS (Murashige & Skoog medium including vitamins and MES buffer, Duchefa. pH of 5.8). The solution was used to inoculate 100 g sterile soil-like matrix. Calcined clay or calcined clay amended with 3% peat were used as matrix to grow plant. Calcined clay was washed twice, autoclaved and dried for several days at +60°C prior to be disposed in sterile magenta boxes. After clay inoculations, surface sterilized seeds were sowed into the matrix. *Arabidopsis thaliana,* Col-0, seeds were surface sterilized with ethanol and stratified overnight in dark at 4°C. Plants were grown for seven weeks in light cabinet at 22°C with 11h light and 54% humidity.

DNA extraction from gnotobiotic growth system. After seven weeks of incubation in light cabinet, plant were harvested and clay were sampled. For plant, shoot and roots were separated.

Shoot were weighted and roots from three to four plants (same magenta box) were pooled and washed with a solution of water containing 1% PBS and 0.02% Silwet L-77, briefly dried with sterile Whatmann glass microfibre filters (GE Healthcare Life Sciences), transferred to Lysing Matrix E (MP Biomedicals), snap-frozen in liquid nitrogen and stored at -80°C. The corresponding clay samples were washed through shaking in 1 % PBS supplemented with 0.02% Silvet L-77, samples were allowed to settle for 15 min at room temperature and then supernatant was collected. Supernatant was afterwards centrifuged for 10 minutes at 4000 rpm in order to pellet the samples. Collected pellet was resuspended in Nuclease-Free water (Qiagen), transferred to Lysing Matrix E (MP Biomedicals), snap-frozen in liquid nitrogen and stored at -80°C. For DNA extractions, samples were homogenized twice by Precellys 24 tissue lyser (Bertin Technologies). DNA was extracted according to the provided protocol from FastDNA SPIN kit for Soil (MP Biomedicals). DNA was eluted in water and stored at -20°C until library preparation for amplicon sequencing.

Library preparation for sequencing V5-V7 regions of 16S rRNA gene. The library was prepared according to two PCR amplification steps. DNA samples were thawed, quantified by fluorimetric method and then concentrations adjusted to ~3.5 ng/µl. The regions V5-V7 of the 16S rRNA genes were amplified during 25 cycles by forward 799F and reverse 1193R primers in 25 µl reaction volume in triplicates. The three PCR reactions were pooled and excess of primers and nucleotides digested by incubating at 37°C with 20U Exonuclease I and 5U Antarctic phosphatase (New England, BioLabs). Enzymes were heat-inactivated by incubating at 85°C for 15 minutes, afterwards samples were centrifuged for 10 minutes at 4000 rpm and supernatant was collected and served as template for the second PCR amplification step. Collected templates were PCR-barcoded with reverse primers (B5-1 to B5-96) that are compatible with Illumina sequencing technologies. All samples were triplicated and amplified during 10 cycles. Replicates from one sample were pooled and run in 1.5% (w/v) agarose gel. Barcoded amplicons were extracted from gel by QIAquick Gel Extraction Kit (Qiagen). Samples were quantified for DNA concentration and pulled at *equi* concentrations. The amplicon libraries were twice cleaned by Agencourt AMPure XP Kit (Beckman Coulter) and submitted for sequencing at an Illumina MiSeq platform using the MiSeq Reagent kit v3 following the 2x 300 bp paired-end sequencing protocol (Illumina Inc.).

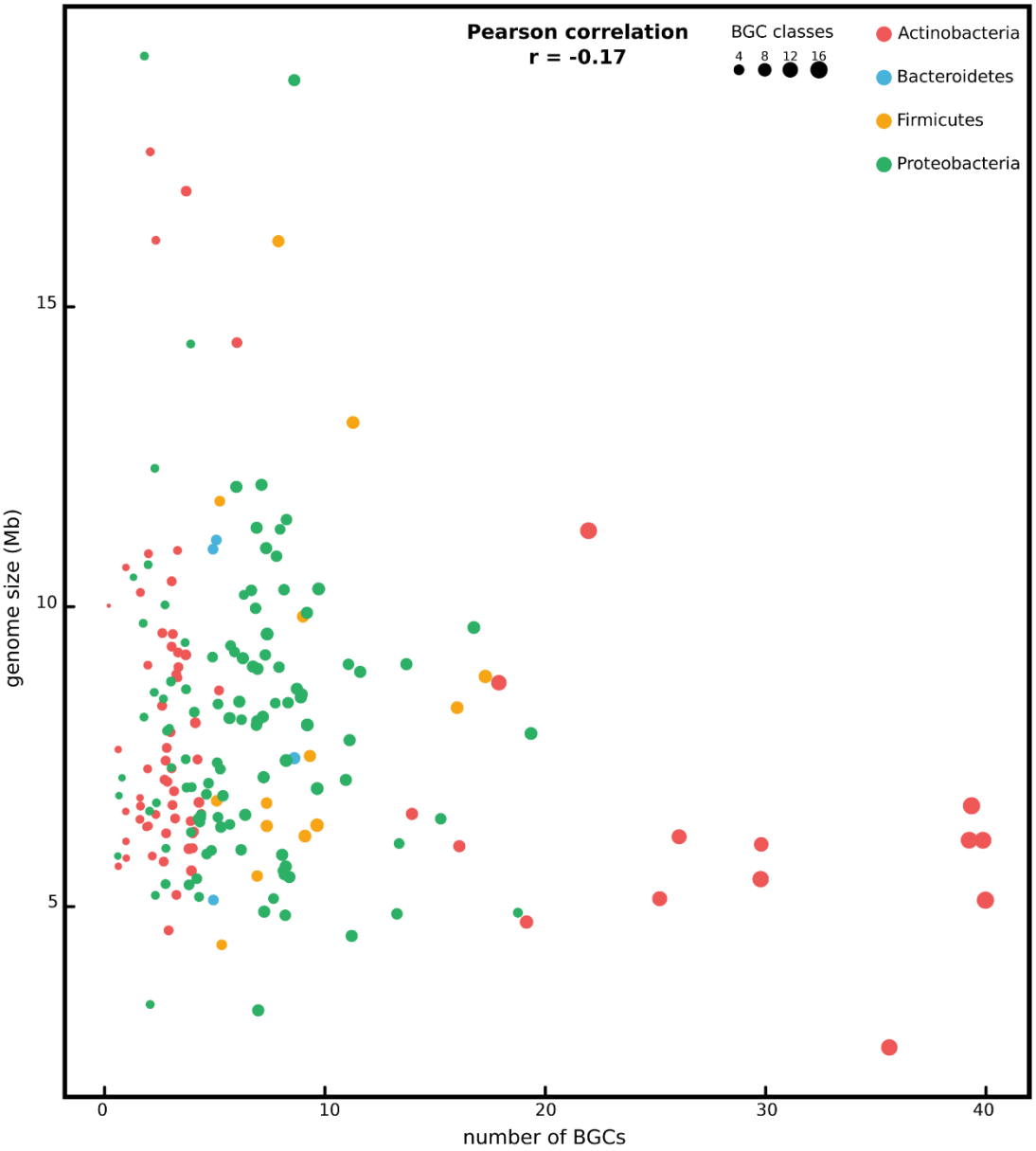
Data analysis of amplicons sequencing. Forward and reverse reads were joined, demultiplex and quality filtered (Phred ≥ 30) using PANDAseq paired-end assembler (Masella et al., 2012).

Afterwards, Pair-end joined reads were split by samples using a Python script (developed by Benli Chai, Michigan State University) and then reads were mapped to reference sequences (V5-V7 regions of the 16S rRNA gene) using the RDP alignment script (Cole *et al.*, 2014). Prior mapping reads to reference sequences, 198 full-length 16S rRNA gene sequences (download from "www.at-sphere.com") were trimmed to V5-V7 regions and de-replicated into 130 sequences at 100% sequence similarity. These 130 trimmed sequences were used as reference to map amplicon reads. Only sequences that mapped at 100% to reference sequences were used to generate species count table. For Alpha-diversity analysis, count reads were rarefied to even sequencing depth based on smallest sample size under R statistical environment (R Development Core Team, 2008) using the R package "phyloseq" (McMurdie and Holmes, 2013). For Beta-diversity analysis, count reads were normalized by cumulative sum scaling normalization factors (Paulson *et al.*, 2013) prior to compute distances between samples. Bray Curtis and weighted UniFrac distances were computed using the R package "phyloseq" (McMurdie and Holmes, 2013). Multivariate dispersion (group homogeneity) for a group of samples was competed using the function Betadisper implemented in R environment under the package "vegan" (Anderson, 2006). Plots were generated using R package "ggplot2" (Wickham, 2009).

References

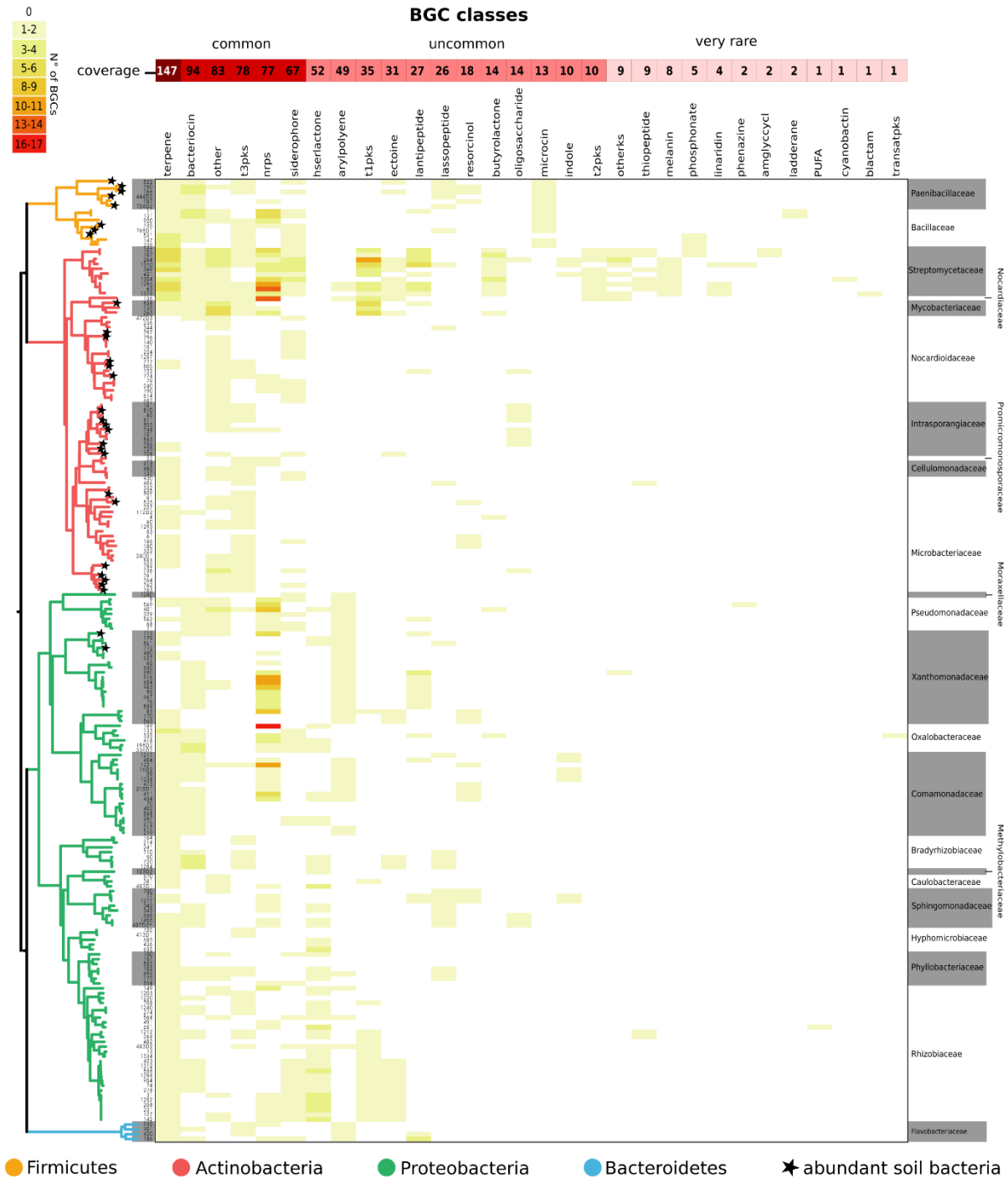
- Anderson, M. J. (2006). Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* 62, 245–253. doi:10.1111/j.1541-0420.2005.00440.x.
- Bai, Y., Müller, D. B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M., et al. (2015). Functional overlap of the Arabidopsis leaf and root microbiota. *Nature* 528, 364– 369. doi:10.1038/nature16192.
- Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., et al. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* 37, D141–D145. doi:10.1093/nar/gkn879.
- Epskamp, S., Cramer, A. O. ., Waldorp, L. J., Schmittmann, V. D., and Borsboom, D. (2012). qgraph: Network Visualizations of Relationships in Psychometric Data. Available at: https://www.jstatsoft.org/article/view/v048i04 [Accessed April 23, 2017].
- Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G., and Neufeld, J. D. (2012). PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics* 13, 31. doi:10.1186/1471-2105-13-31.
- McMurdie, P. J., and Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* 8, e61217. doi:10.1371/journal.pone.0061217.
- Paulson, J. N., Stine, O. C., Bravo, H. C., and Pop, M. (2013). Differential abundance analysis for microbial marker-gene surveys. *Nat Meth* 10, 1200–1202. doi:10.1038/nmeth.2658.
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Meth* 9, 671–675. doi:10.1038/nmeth.2089.
- Wickham, H. (2009). *ggplot2 - Elegant Graphics for Data Analysis | Hadley Wickham | Springer*. Available at: http://www.springer.com/br/book/9780387981413 [Accessed April 23, 2017].

Supplementary Figures and Tables



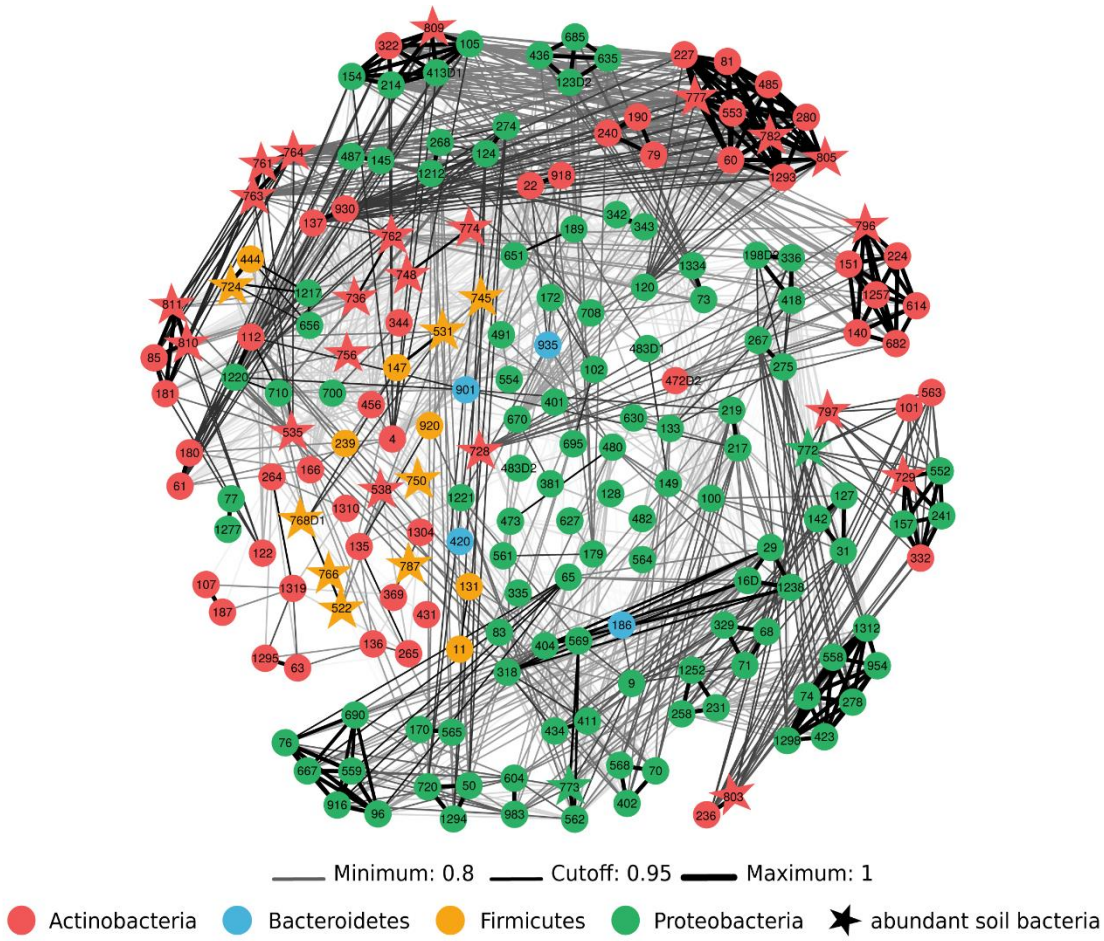
Supplementary-figure 1| **The genome size do not correlate with the number of BGCs.**

The plot shows Pearson correlation between the genome size and the number of predicted BGCs. Each shape indicates an isolate and the size of the shape indicates the number of BGC classes.



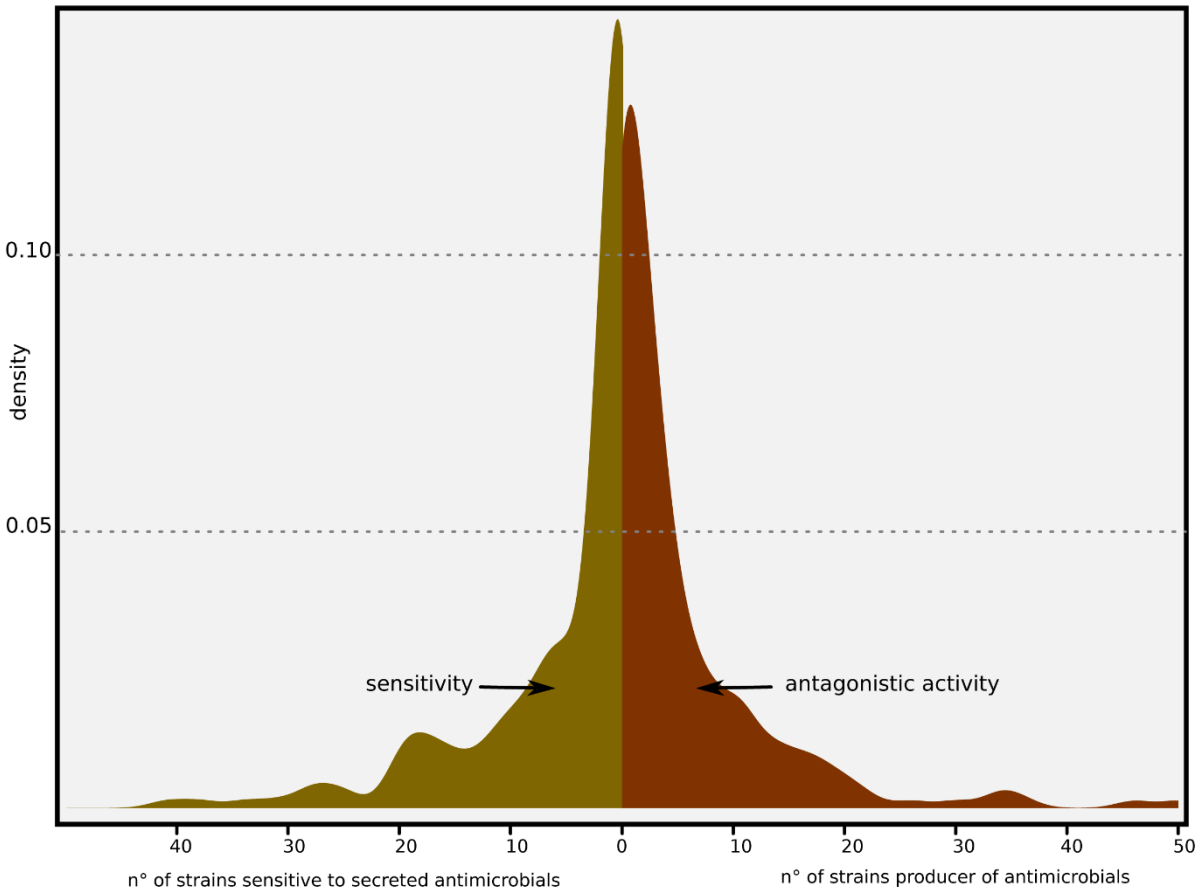
Supplementary-figure 2| **The terpene class is widely predicted across the bacterial strains.**

The heatmap shows percentage of BGCs across all the 198 bacterial isolates. Color code inside the heatmap indicates the enrichement of BGC class compared to other class for one isolate. The bacterial isolates are depicted in the y-axis and abundant soil bacteria are indicated by stars. BGC classes are indicated in the x-axis. Most of BGC classes are very rare to uncommon. The terpene class is widely prideicted from the 198 bacterial genomes.



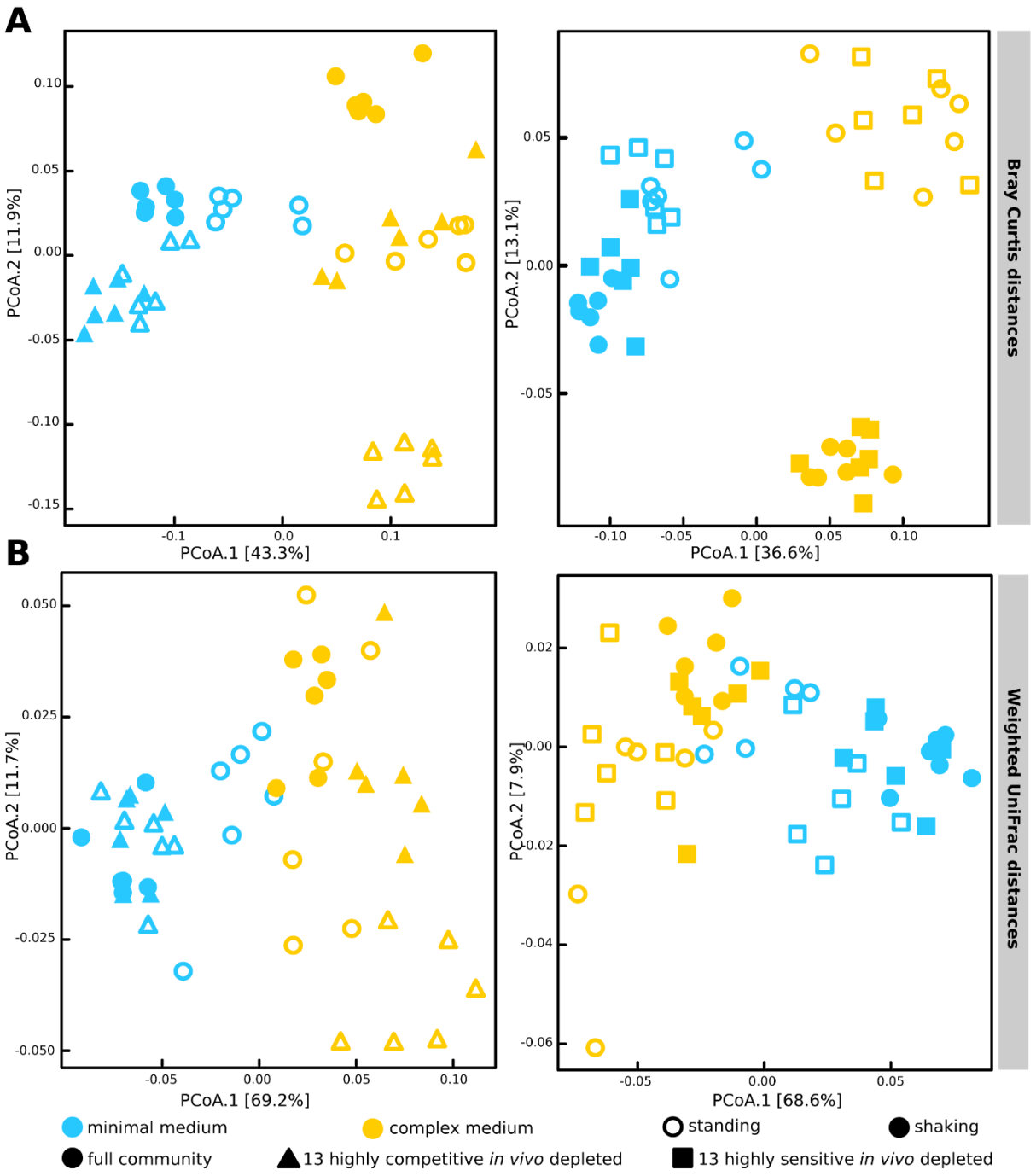
Supplementary-figure 3| **Correlation network of the BGCs profiles.**

The graph shows correlation network of the BGCs profiles of root-derived and abundant soil bacteria. Nodes represent bacterial isolates and are colored coded according to phylum. Abundant soil bacteria are represented by stars. Edges show positive correlation coefficient. The correlation coefficients are above 0.8 for the significance level of an FDR-corrected *p-value* of 0.05. The color gradient in the edges from bright to dark and thickness from thin to large indicate the degree of correlation from low to high. The bacterial strains show rather diffirent pattern of BGC classes.



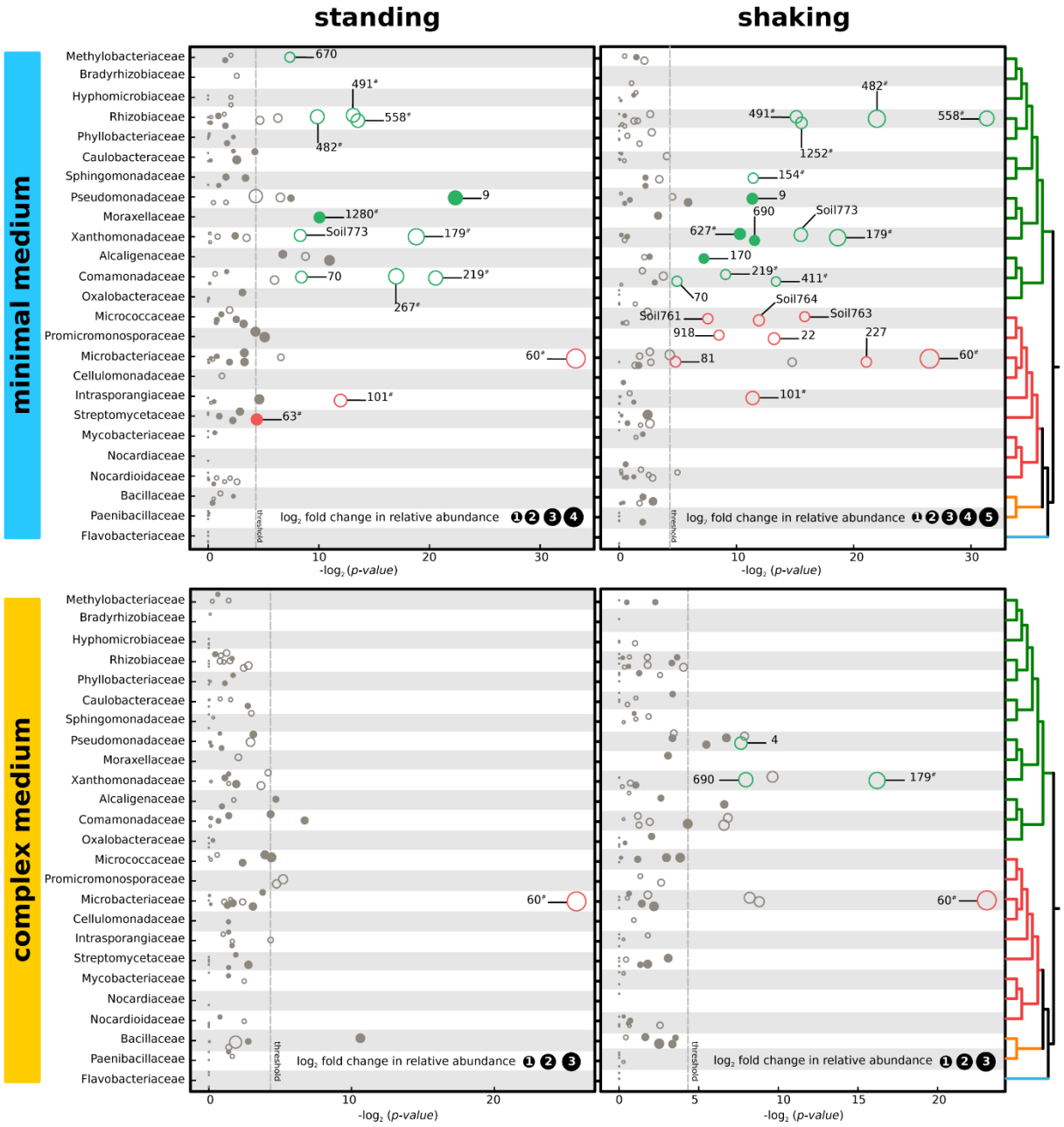
Supplementary-figure 4| **Sensitivity and antagonistic activity density plots.**

The sensitivity density plot shows that most of the isolates are inhibited between 1 to 10 times. The antagonistic activity density plot indicates that bacteria inhibit between 1 to 10 other bacterial strains. Only fewer bacterial isolates inhibit or are inhibited by more that 20 other bacterial strains.



Supplementary-figure 5| **Shift in community structure upon perturbation by community members' depletion.**

Panel-A and **-B** depict principle coordinates analysis on Bray Curtis (BC) and weighted UniFrac (wUF) distances between samples, receptively. Circles represent full community *in silico* depleted from highly competitive bacteria in the left-side plot or full community *in silico* depleted from highly sensitive bacteria in right-side plot in both panels. Triangles refer to the 13 highly competitive isolates *in vivo* depleted communities. Square shapes refer to the 13 highly sensitive isolates *in vivo* depleted communities. Filled shapes indicate shaking state and unfilled shapes indicate standing state. Color code in shapes indicate minimal or complex medium. **Panel-A**, PCoA on BC distances indicate that growth medium and growth state explain most of observed variance. **Panel-B**, same conclusion hold true based on wUF distances. Based on both distance metrics, communities *in vivo* depleted from the 13 highly competitive strains tend to cluster far from corresponding full communities than communities *in vivo* depleted from the 13 highly sensitive strains.



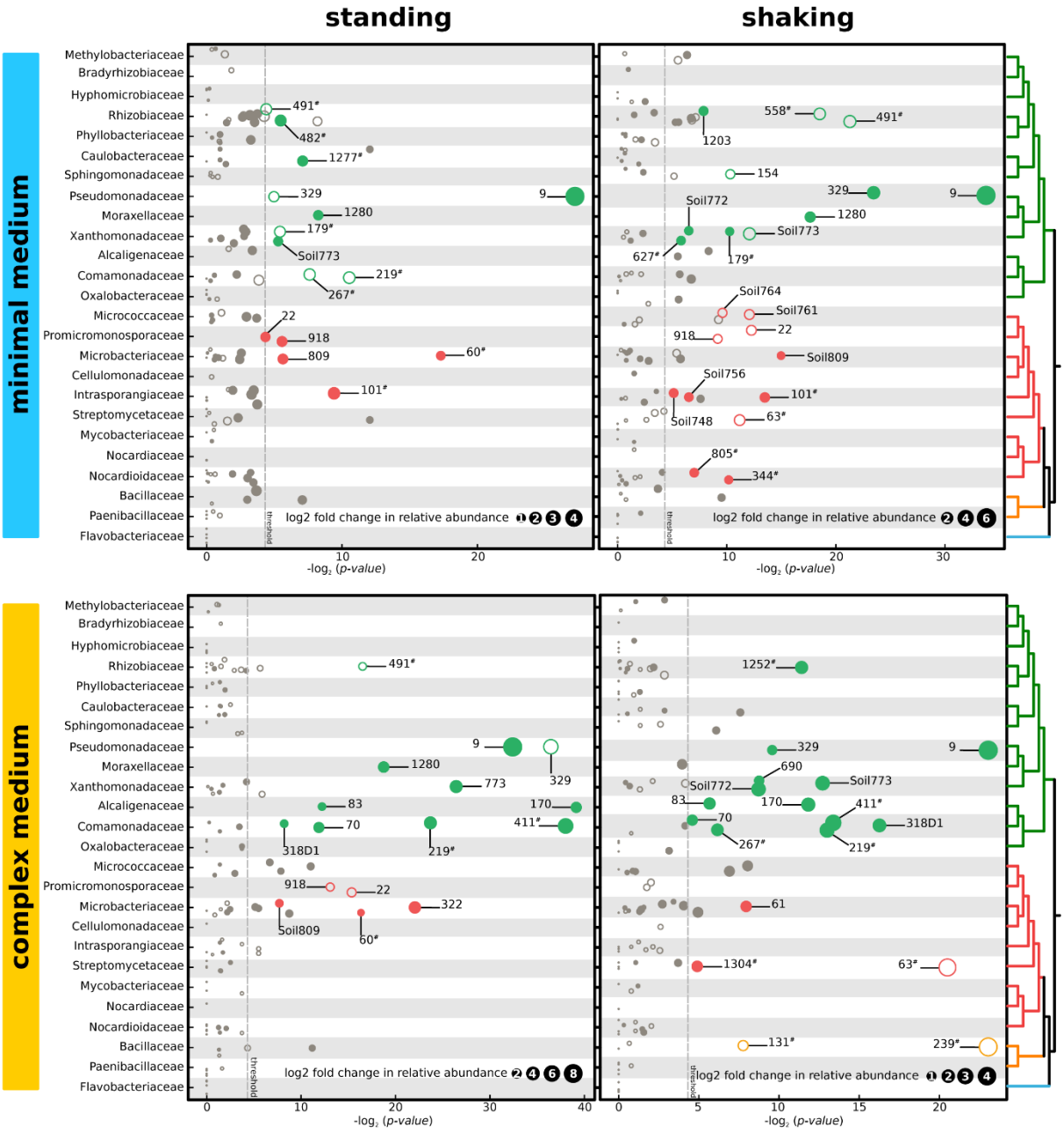
threshold: fdr-corrected *p-value* < 0.05 and $\log_2$ fold change in the relative abundance > 1

● significant increase in relative abundance ○ significant decrease in relative abundance

● Proteobacteria ● Actinobacteria ● Firmicutes # representative of mutiple isolates * abundant soil bacteria

Supplementary-figure 6| ***In vivo* depletion of the 13 highly sensitive bacteria leads mainly to a decrease in relative abundance of several bacterial isolates.**

The plots show changes in the relative abundance of bacterial isolates in output communities after *in vivo* depletion of the 13 highly sensitive bacteria. Each circle represent an isolate. Color code indicates phylum. Filled circle indicates a significant increase in the relative abundance and unfilled circle indicates a significant decrease in the relative abundance. Bacterial families are indicated in y-axis and x-axis indicates $\log_2$ transformed FDR-corrected *p-value* after the enrichment test. The size of circles indicates $\log_2$ fold change in the relative abundance.



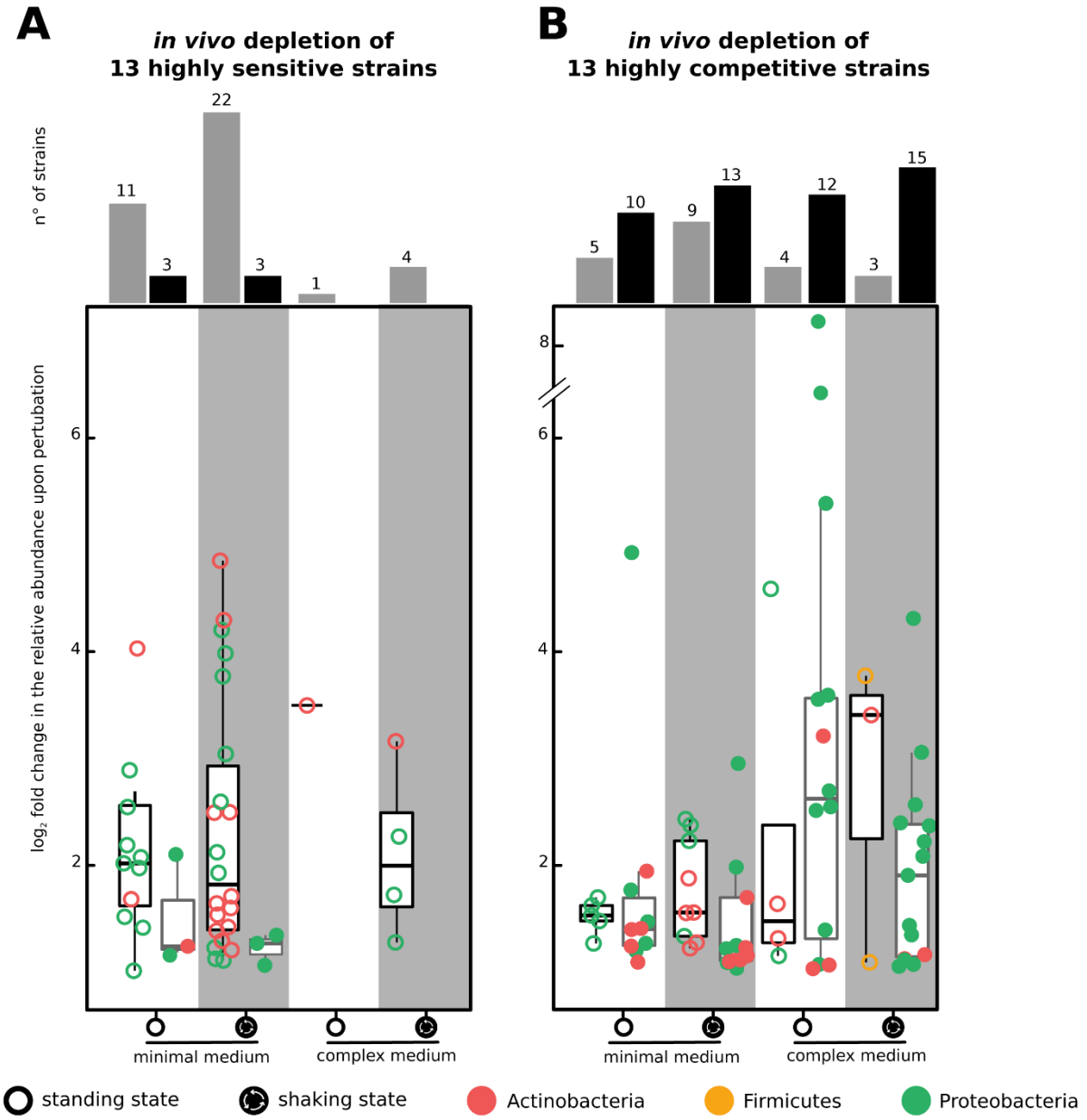
threshold: *p-value* < 0.05 and $\log_2$ fold change in the relative abundance > 1

● significant increase in relative abundance ○ significant decrease in relative abundance

● Proteobacteria ● Actinobacteria ● Firmicutes # representative of mutiple isolates *: abundant soil bacteria

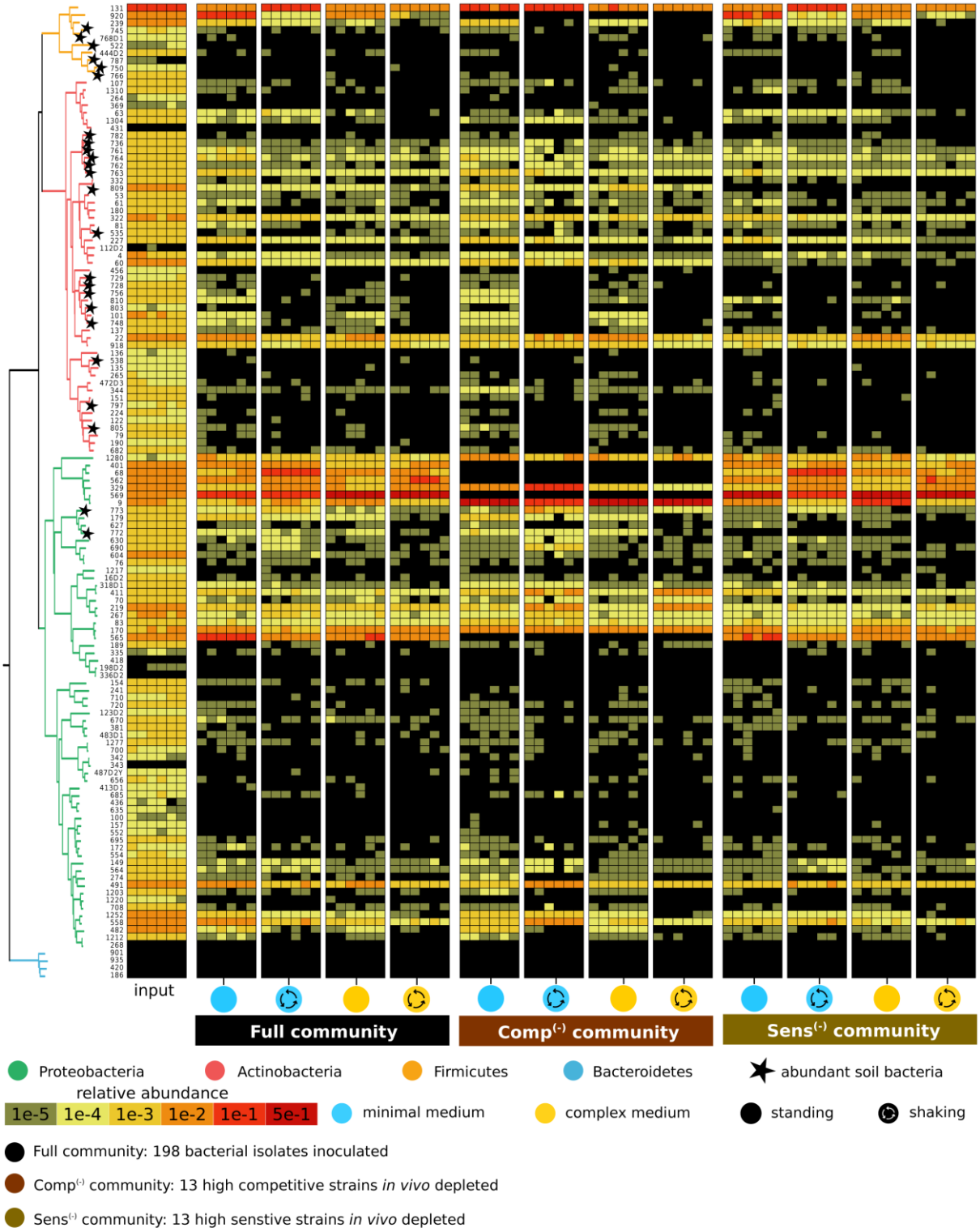
Supplementary-figure 7| ***In vivo* depletion of the 13 highly competitive strains leads to the enrichment of several bacterial isolates.**

The plots show $\log_2$ fold change in the relative abundance of the bacterial isolates in output communities after *in vivo* depletion of the 13 highly competitive strains. Each circle represent an isolate. Color code indicates phylum. Filled circle indicates a significant increase in the relative abundance and unfilled circle indicates a significant decrease in the relative abundance. Bacterial families are indicated in y-axis and x-axis indicates $\log_2$ transformed FDR-corrected *p-value* of the enrichment test. Size in circles indicates $\log_2$ fold change in the relative abundance.



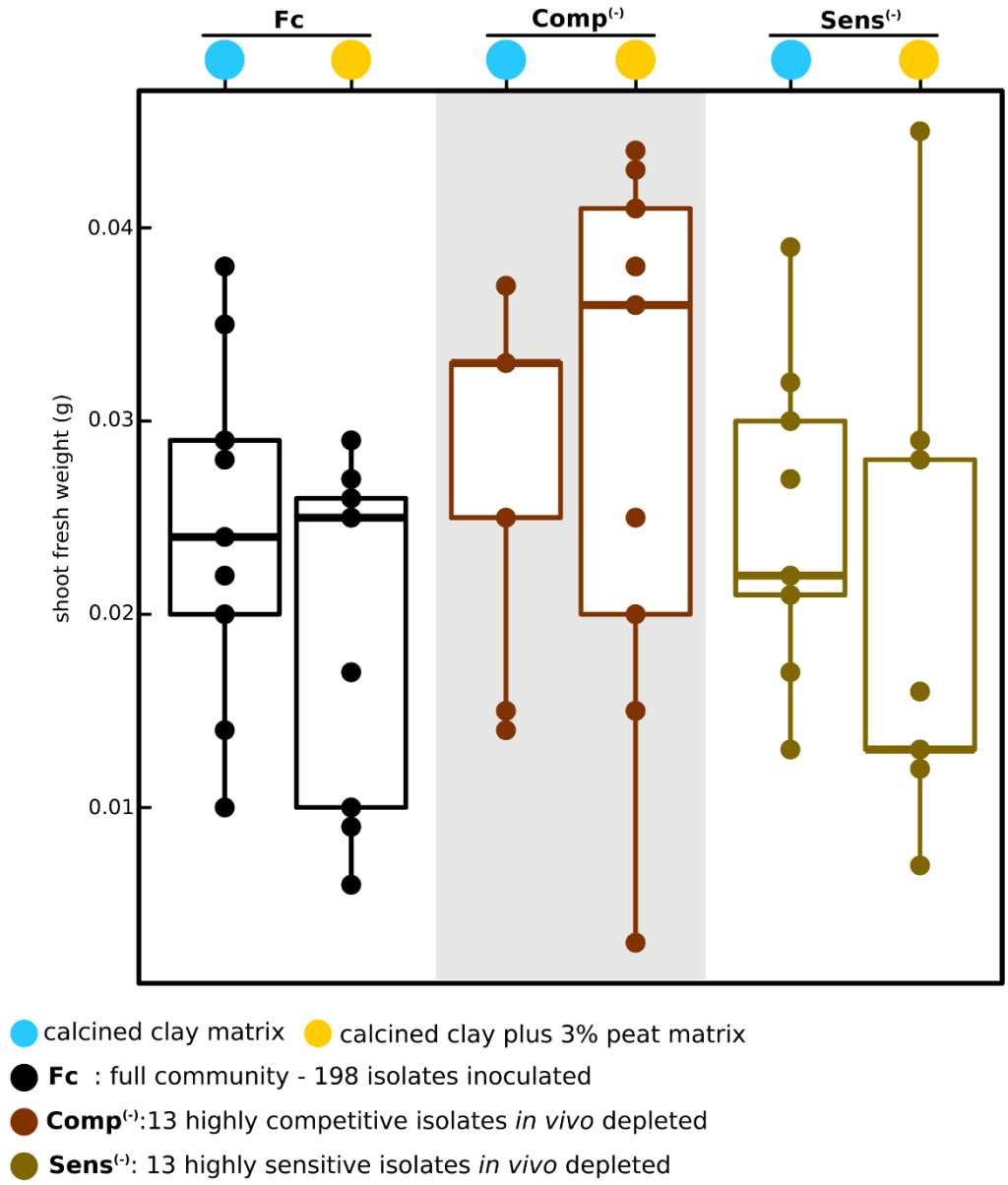
Supplementary-figure 8| **Significant increase and decrease in the relative abundance of bacterial isolates upon perturbation.**

Panel-A and **-B** depicting $\log_2$ fold change in the relative abundance of the bacterial isolates that significantly increase or decrease upon community members' depletions, respectively. Filled circle indicates a significant increase in the relative abundance and unfilled circle indicates a significant decrease in the relative abundance in perturbed communities compared to the corresponding full communities *in silico* depleted from highly sensitive or competitive bacteria. Each circle represent an isolate. Bar plots indicate total number of isolates that significantly increase or decrease in the relative abundance in each microcosm. **Panel-A**, *in vivo* depletion of the 13 highly sensitive strains leads mainly to a significant decrease in the relative abundance of the bacterial isolates. **Panel-B**, *in vivo* depletion of the 13 highly competitive strains leads to a significant increase in the relative abundance of the bacteria.



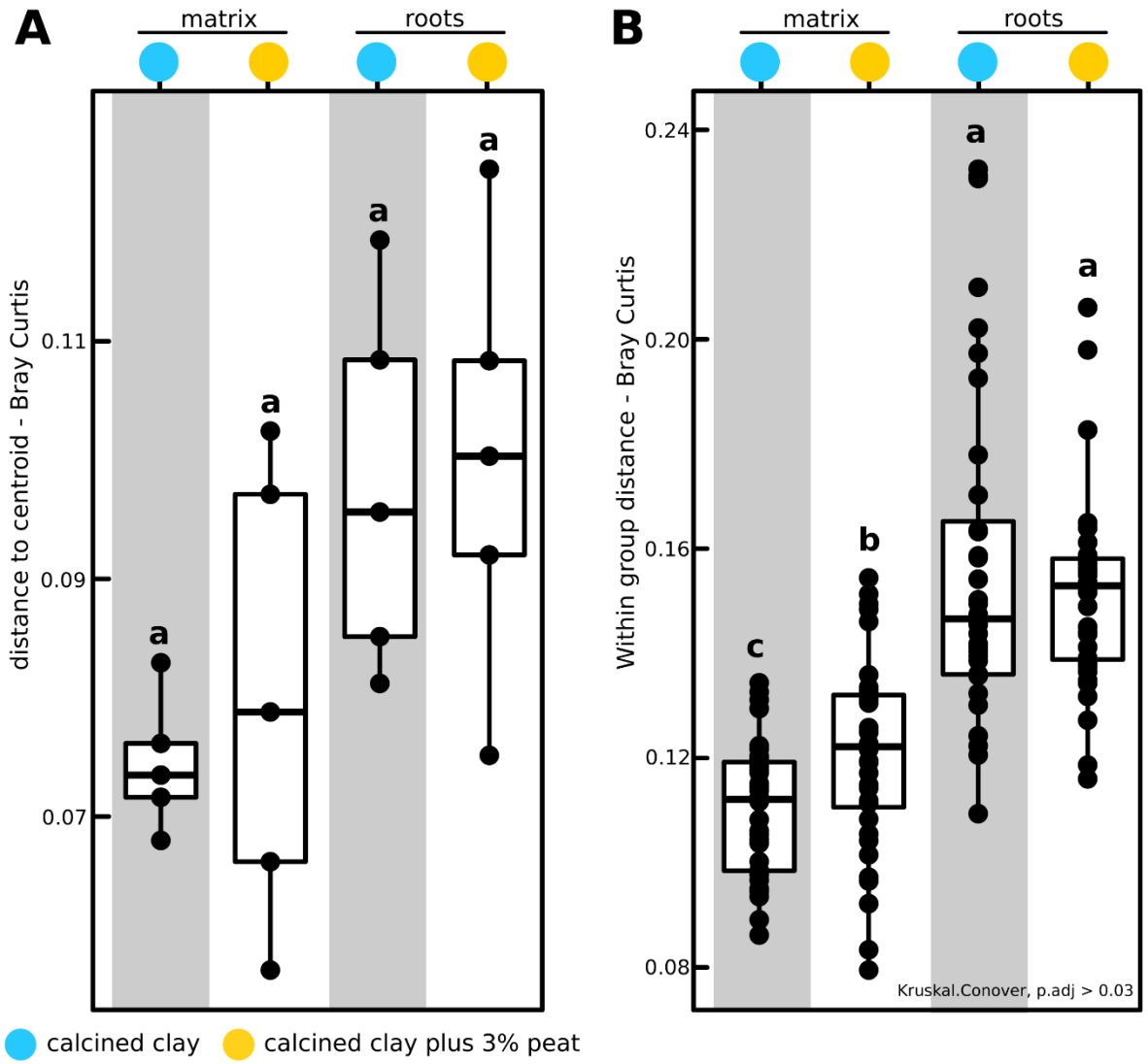
Supplementary-figure 9| **Relative abundance of the bacterial isolates in input and output communities.**

The heatmap shows the relative abundance of input and the output communities. Output communities are full communities, the 13 highly competitive strains *in vivo* depleted communities, or the 13 highly sensitive strains *in vivo* depleted communities. Color code in the tree indicates phylum and star indicates abundant soil bacteria. Each microcosm is represented by three biological replicates and two technical replicates.



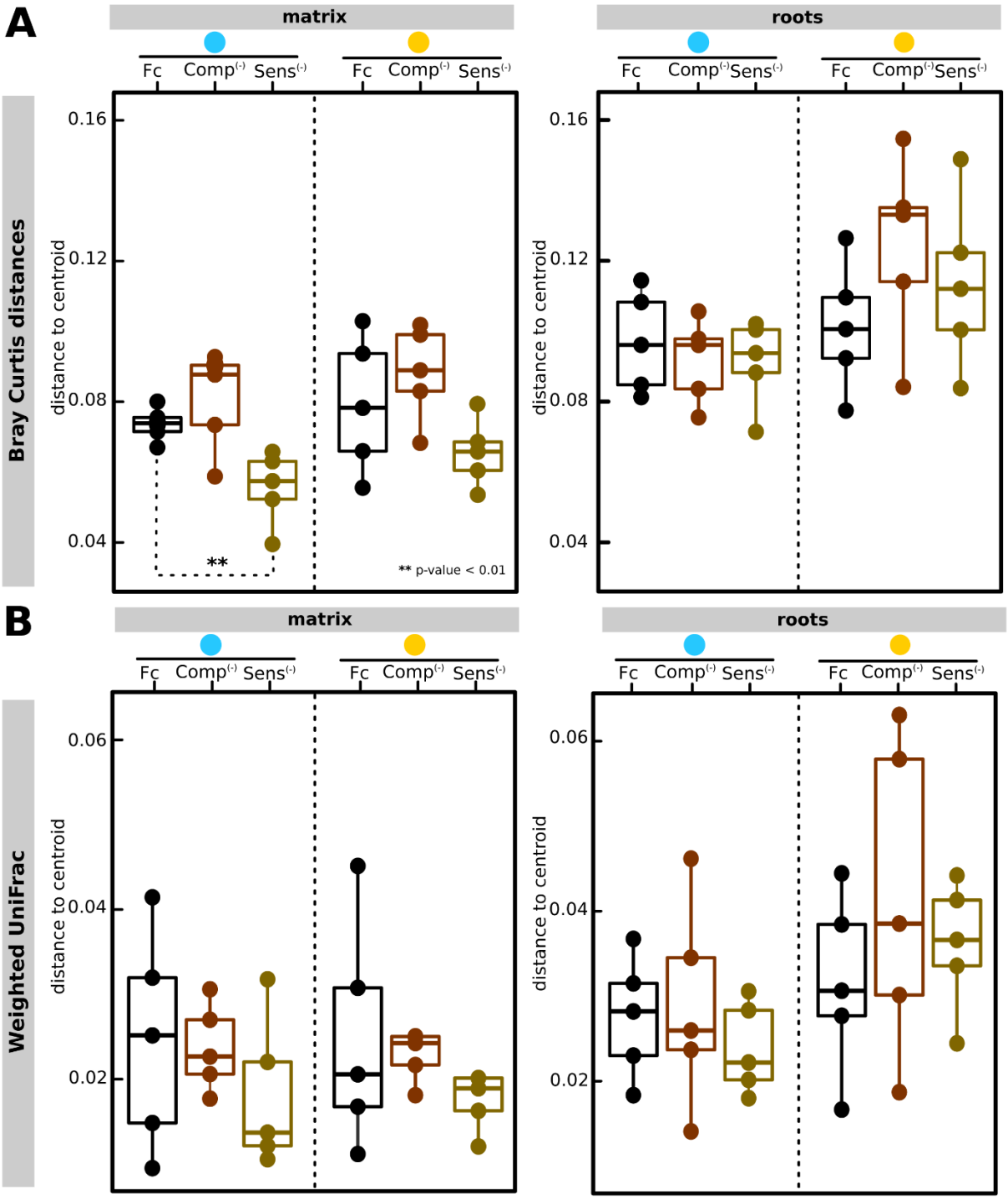
Supplementary-figure 10| **Shoot fresh weight is not affected by *in vivo* depletion of community members'.**

The box plots depicts fresh shoot weight of *A. thaliana* after 7 weeks of growth in phytochamber. Each circle correspond to a mean of three to four plants. Color code indicates full community (198 bacterial isolates), *in vivo* depleted from the 13 highly competitive strains communities, or *in vivo* depleted from the 13 highly sensitive strains communities.



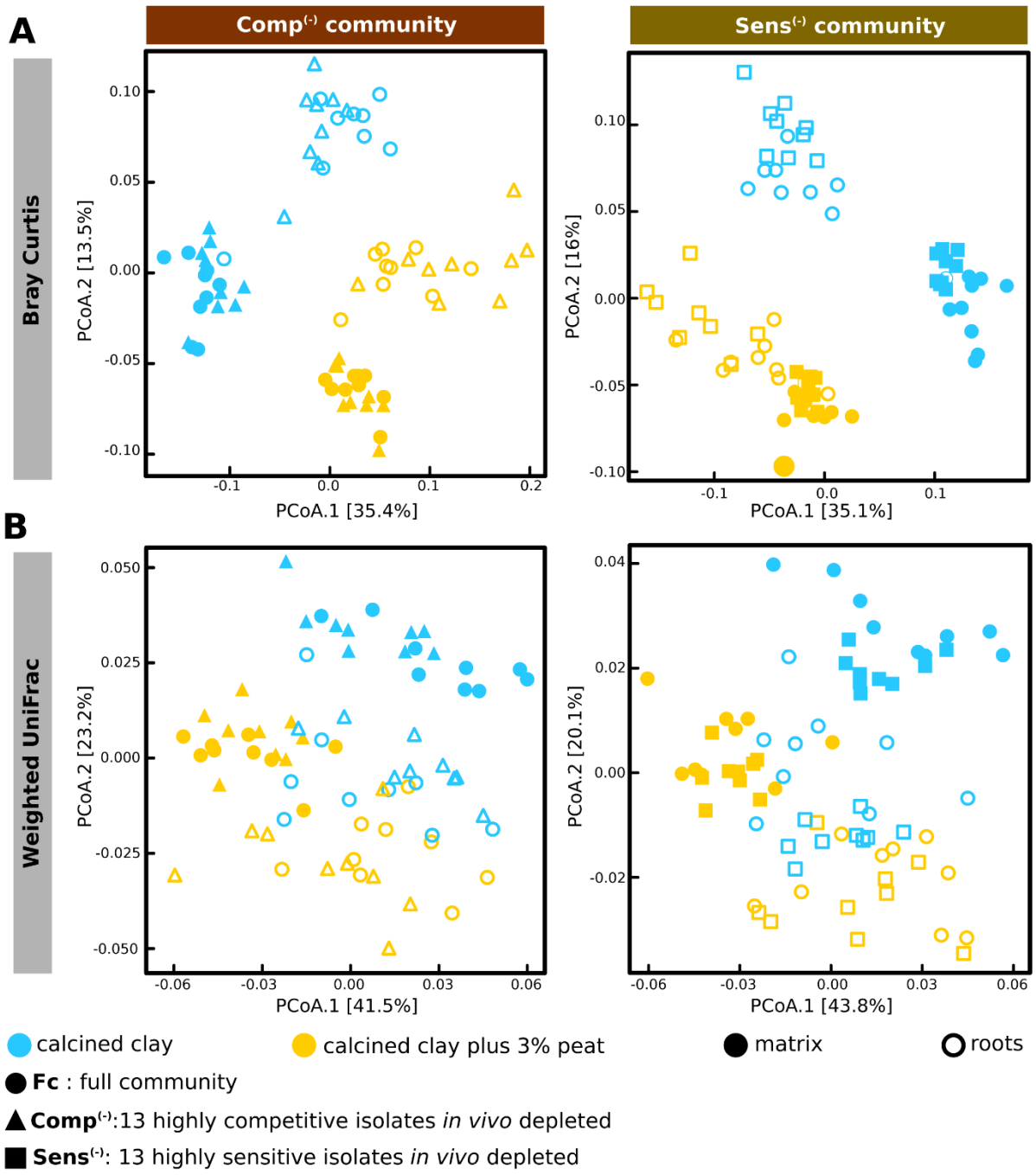
Supplementary-figure 11| **The bacterial communities in the matrix are more homogeneous than in the roots.**

Panel-A and **-B** depict Bray Curtis (BC) distance to centroid and within group (BC) distance, respectively. **Panel-A**, although there is no significant difference in BC distance to centroid between groups, the bacterial communities in the roots for both calcined clay and calcined clay plus 3% peat systems show high mean values in the distance to centroid. **Panel-B**, matrix bacterial communities show significantly low within group BC distance than the roots bacterial communities. The bacterial communities in calcined clay matrix are significantly more homogeneous than the bacterial in calcined clay plus 3% peat.



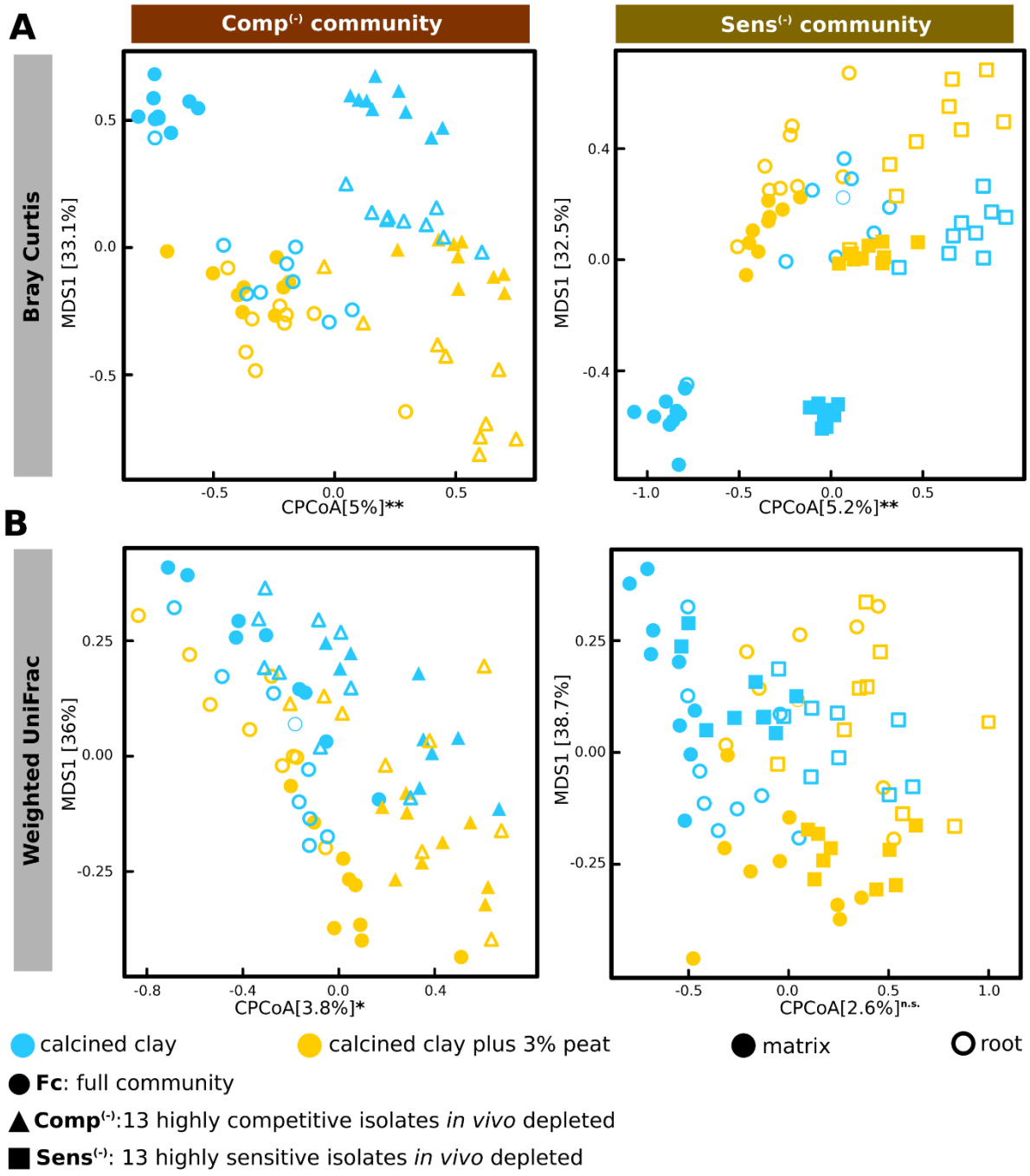
Supplementary-figure 11| **Both community members' depletions do not alter significantly the homogeneity of output communities.**

Panel-A and **-B** depict Bray Curtis "BC" and weighted UniFrac "wUF" distance to centroid, respectively. Each circle indicates an output bacterial community and color code indicates full community, Comp$^{(-)}$ community and Sens$^{(-)}$ community. **Panel-A**, Perturbation by community members depletion do not alter significantly the homogeneity of the bacterial communities, except *in vivo* depletion of the 13 highly sensitive strains in calcined clay matrix. **Panel-B**, similar analysis based on based on WuF distances indicates also that group homogeneity in outputs communities is not significantly altered.



Supplementary-figure 12| **Principal coordinates analysis of perturbed and unperturbed output communities.**

Panel-A and **-B** depict principle coordinate analysis on Bray Curtis (BC) and weighted UniFrac (wUF) distances between samples, receptively. Circles represent full community *in silico* depleted from highly competitive bacteria in the left-side plots or full community *in silico* depleted from highly sensitive bacteria in right-side plots, for both panels. Triangles refer to communities *in vivo* depleted from the 13 highly competitive strians and square refer to communities *in vivo* depleted from the 13 highly sensitive strains. Filled shapes indicate matrix and unfilled shapes indicate roots. Color code in shapes indicate calcined clay or calcined clay plus 3% peat. **Panel-A**, PCoA on BC distances shows that habitat (matrix, roots) and 3% peat amendment explain most of the observed variance. **Panel-B**, PCoA on wUF the observation of the panel-A.



Supplementary-figure 13| **Both *in vivo* community members' depletions show a comparable percetnage of observed variance in output communities.**

Panel-A and **-B** depict constrained principle coordinate analysis of Bray Curtis (BC) and weighted UniFrac (wUF) distances, receptively. PCoA plots are constrained by *in vivo* depletion of the 13 highly competitive bacteria in left-side plot or by in vivo depletion of 13 highly sensitive bacteria in right-side plot for both panels. **Panel-A**, *in vivo* depletion of the 13 highly competitive bacteria explains 5% of the observed variance and *in vivo* depletion of the 13 highly sensitive bacteria explains 5.2% of the observed variance, both observed shifts in the community structure are significant (p.value <0.01). **Panel-B**, constrained coordinates analysis on wUF distances indicates that both *in vivo* depletions explain less variance (below 5%) with *p-value* >0.01.

Supplementary table 1l **Taxonomy of the bacterial isolates.**

Current table shows taxonomy of 198 bacterial isolates employed in exploring the genomes for biosynthetic genes clusters, analysis of metabolites, screen for inter-bacterial antagonistic interactions and *in vitro* and *in planta* perturbation experiments.

ID	group	phylum	class	order	family	Genus
4	Root	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	Agromyces
9	Root	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas
11	Root	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus
22	Root	Actinobacteria	Actinobacteria	Actinomycetales	Promicromonosporaceae	ND
29	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	ND
31	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Sinorhizobium
50	Root	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas
53	Root	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	Microbacterium
60	Root	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	ND
61	Root	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	Microbacterium
63	Root	Actinobacteria	Actinobacteria	Actinomycetales	Streptomycetaceae	Streptomyces
65	Root	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Pseudoxanthomonas
68	Root	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas
70	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax
71	Root	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas
73	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium
74	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Sinorhizobium
76	Root	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	ND
77	Root	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	ND
79	Root	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Nocardioides
81	Root	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	Agromyces
83	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Alcaligenaceae	Achromobacter
85	Root	Actinobacteria	Actinobacteria	Actinomycetales	Intrasporangiaceae	ND
96	Root	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	ND
100	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	ND
101	Root	Actinobacteria	Actinobacteria	Actinomycetales	Intrasporangiaceae	Janibacter
102	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	Mesorhizobium
105	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	ND
107	Root	Actinobacteria	Actinobacteria	Actinomycetales	Streptomycetaceae	ND
122	Root	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Nocardioides
127	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Sinorhizobium
131	Root	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus

133	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Massilia
135	Root	Actinobacteria	Actinobacteria	Actinomycetales	Mycobacteriaceae	Mycobacterium
136	Root	Actinobacteria	Actinobacteria	Actinomycetales	Nocardiaceae	Nocardia
137	Root	Actinobacteria	Actinobacteria	Actinomycetales	Cellulomonadaceae	Cellulomonas
140	Root	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Nocardioides
142	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Sinorhizobium
147	Root	Firmicutes	Bacilli	Bacillales	Bacillaceae	ND
149	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium
151	Root	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Nocardioides
154	Root	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	ND
157	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	Mesorhizobium
166	Root	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	Microbacterium
170	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Alcaligenaceae	Achromobacter
172	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	Mesorhizobium
179	Root	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Rhodanobacter
180	Root	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	ND
181	Root	Actinobacteria	Actinobacteria	Actinomycetales	Intrasporangiaceae	ND
186	Root	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
187	Root	Actinobacteria	Actinobacteria	Actinomycetales	Streptomycetaceae	ND
189	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Herbaspirillum
190	Root	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Nocardioides
214	Root	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingopyxis
217	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax
219	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax
224	Root	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Nocardioides
227	Root	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	ND
231	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Sinorhizobium
236	Root	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Aeromicrobium
239	Root	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus
240	Root	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Nocardioides
241	Root	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas
258	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Sinorhizobium
264	Root	Actinobacteria	Actinobacteria	Actinomycetales	Streptomycetaceae	Streptomyces
265	Root	Actinobacteria	Actinobacteria	Actinomycetales	Mycobacteriaceae	Mycobacterium
267	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax
268	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium
274	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	ND
275	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax

278	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Sinorhizobium
322	Root	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	Microbacterium
329	Root	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas
332	Root	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	ND
335	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Massilia
342	Root	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Caulobacter
343	Root	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Caulobacter
344	Root	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Aeromicrobium
369	Root	Actinobacteria	Actinobacteria	Actinomycetales	Streptomycetaceae	Streptomyces
381	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylobacteriaceae	ND
401	Root	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas
402	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Acidovorax
404	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	ND
411	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Variovorax
418	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Janthinobacterium
420	Root	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
423	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Sinorhizobium
431	Root	Actinobacteria	Actinobacteria	Actinomycetales	Streptomycetaceae	Streptomyces
434	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Variovorax
436	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	ND
456	Root	Actinobacteria	Actinobacteria	Actinomycetales	Intrasporangiaceae	ND
473	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Variovorax
480	Root	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	ND
482	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium
485	Root	Actinobacteria	Actinobacteria	Actinomycetales	Cellulomonadaceae	Cellulomonas
491	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Agrobacterium
522	Soil	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Paenibacillus
531	Soil	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus
535	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	Agromyces
538	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Mycobacteriaceae	Mycobacterium
552	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	Mesorhizobium
553	Root	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	ND
554	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	Mesorhizobium
558	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Sinorhizobium
559	Root	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	ND
561	Root	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Rhodanobacter
562	Root	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas
563	Root	Actinobacteria	Actinobacteria	Actinomycetales	Intrasporangiaceae	Janibacter

564	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Agrobacterium
565	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Alcaligenaceae	Achromobacter
568	Root	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	ND
569	Root	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas
604	Root	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	ND
614	Root	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	ND
627	Root	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Rhodanobacter
630	Root	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Pseudoxanthomonas
635	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	ND
651	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Agrobacterium
656	Root	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Caulobacter
667	Root	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	ND
670	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylobacteriaceae	ND
682	Root	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	ND
685	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Hyphomicrobiaceae	ND
690	Root	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	ND
695	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	Mesorhizobium
700	Root	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	ND
708	Root	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium
710	Root	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas
720	Root	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas
728	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Intrasporangiaceae	Janibacter
729	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Intrasporangiaceae	ND
736	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Micrococcaceae	Arthrobacter
745	Soil	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus
748	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Intrasporangiaceae	Janibacter
750	Soil	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Paenibacillus
756	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Intrasporangiaceae	Janibacter
761	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Micrococcaceae	Arthrobacter
762	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Micrococcaceae	Arthrobacter
763	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Micrococcaceae	Arthrobacter
764	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Micrococcaceae	Arthrobacter
766	Soil	Firmicutes	Bacilli	Bacillales	Paenibacillaceae	Paenibacillus
772	Soil	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Rhodanobacter
773	Soil	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Rhodanobacter
774	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Nocardioides
777	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Nocardioides
782	Soil	Actinobacteria	Actinobacteria	Actinomycetales	Micrococcaceae	Arthrobacter

787	*Soil*	*Firmicutes*	*Bacilli*	*Bacillales*	*Paenibacillaceae*	*Paenibacillus*
796	*Soil*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Nocardioidaceae*	*Nocardioides*
797	*Soil*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Nocardioidaceae*	*Nocardioides*
803	*Soil*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Intrasporangiaceae*	*Janibacter*
805	*Soil*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Nocardioidaceae*	*Nocardioides*
809	*Soil*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Microbacteriaceae*	*ND*
810	*Soil*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Intrasporangiaceae*	*ND*
811	*Soil*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Intrasporangiaceae*	*ND*
901	*Root*	*Bacteroidetes*	*Flavobacteriia*	*Flavobacteriales*	*Flavobacteriaceae*	*Flavobacterium*
916	*Root*	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*ND*
918	*Root*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Promicromonosporaceae*	*ND*
920	*Root*	*Firmicutes*	*Bacilli*	*Bacillales*	*Bacillaceae*	*Bacillus*
930	*Root*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Cellulomonadaceae*	*ND*
935	*Root*	*Bacteroidetes*	*Flavobacteriia*	*Flavobacteriales*	*Flavobacteriaceae*	*Flavobacterium*
954	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Rhizobium*
983	*Root*	*Proteobacteria*	*Gammaproteobacteria*	*Xanthomonadales*	*Xanthomonadaceae*	*ND*
1203	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Rhizobium*
1212	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Rhizobium*
1217	*Root*	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	*ND*
1220	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Rhizobium*
1221	*Root*	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	*ND*
1238	*Root*	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	*ND*
1240	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Agrobacterium*
1252	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Sinorhizobium*
1257	*Root*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Nocardioidaceae*	*Nocardioides*
1277	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Caulobacterales*	*Caulobacteraceae*	*ND*
1280	*Root*	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Moraxellaceae*	*Acinetobacter*
1293	*Root*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Microbacteriaceae*	*ND*
1294	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Sphingomonas*
1295	*Root*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Streptomycetaceae*	*Streptomyces*
1298	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Rhizobium*
1304	*Root*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Streptomycetaceae*	*Streptomyces*
1310	*Root*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Streptomycetaceae*	*Streptomyces*
1312	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Sinorhizobium*
1319	*Root*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Streptomycetaceae*	*Streptomyces*
1334	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Rhizobium*
1455	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Caulobacterales*	*Caulobacteraceae*	*Caulobacter*
112D2	*Root*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Microbacteriaceae*	*ND*

123D2	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Bradyrhizobiaceae*	*Afipia*
16D2	*Root*	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	*ND*
198D2	*Root*	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Oxalobacteraceae*	*ND*
280D1	*Root*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Microbacteriaceae*	*Microbacterium*
318D1	*Root*	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	*Variovorax*
336D2	*Root*	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Oxalobacteraceae*	*ND*
413D1	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Hyphomicrobiaceae*	*ND*
444D2	*Root*	*Firmicutes*	*Bacilli*	*Bacillales*	*Paenibacillaceae*	*Paenibacillus*
472D3	*Root*	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Nocardioidaceae*	*ND*
483D1	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Methylobacteriaceae*	*Methylobacterium*
483D2	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Rhizobiaceae*	*Rhizobium*
487D2Y	*Root*	*Proteobacteria*	*Alphaproteobacteria*	*Caulobacterales*	*Caulobacteraceae*	*Caulobacter*
724D2	*Soil*	*Firmicutes*	*Bacilli*	*Bacillales*	*Paenibacillaceae*	*Paenibacillus*
768D1	*Soil*	*Firmicutes*	*Bacilli*	*Bacillales*	*Bacillaceae*	*Bacillus*

Supplementary Table 2l **Number of biosynthetic genes clusters in bacterial genomes.**

Displayed table indicates number of biosynthetic gene clusters across bacterial genomes. Isolate are indicated by their identification number in first column and BGC classes are indicated in first row.

	siderophore	bacteriocin	nrps	terpene	PKS	arylpolyene	butyrolactone	ectoine	hserlactone	oligosaccharide	lantipeptide	resorcinol	PUFA	lassopeptide	microcin	melanin	indole	phosphonate	thiopeptide	cyanobactin	blactam	phenazine	amglyccycl	ladderane	linaridin
4	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	1	2	4	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	1	4	6	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	2	0
22	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	2	2	2	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
31	1	0	1	1	1	0	0	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	0	3	0	1	1	0	0	1	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
60	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
61	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
63	2	2	14	6	5	1	1	1	0	0	3	0	0	0	0	2	0	0	1	0	0	0	0	0	1
65	0	2	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
68	1	2	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
70	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
71	1	2	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
73	0	0	0	1	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
74	1	1	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
76	0	2	4	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
77	0	0	1	1	0	0	0	0	0	0	1	1	0	2	0	0	1	0	0	0	0	0	0	0	0
79	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
81	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
83	0	0	8	2	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
85	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
96	0	2	4	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
100	1	0	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
101	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
102	0	1	0	1	1	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
105	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
107	2	2	6	6	8	0	1	0	0	0	4	0	0	0	0	0	0	1	1	0	0	0	1	0	0
122	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
127	1	0	1	1	1	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
131	1	4	6	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	0
133	1	2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
135	0	1	2	2	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
136	0	1	14	4	7	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
137	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
140	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

142	1	0	1	1	1	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
147	1	2	0	3	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0
149	0	0	3	1	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
151	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
154	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
157	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
166	1	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
170	0	0	0	1	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
172	0	1	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
179	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
180	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
181	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
186	1	1	1	2	0	1	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
187	2	2	3	5	5	0	3	0	0	0	4	0	0	0	0	0	0	1	1	0	0	0	1	0	0
189	0	0	17	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
190	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
214	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
217	0	2	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
219	0	2	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
224	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
227	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
231	1	0	1	1	1	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
236	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
239	1	0	0	3	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0
240	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
241	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
258	1	0	1	1	1	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
264	3	2	2	5	15	0	1	1	0	1	3	0	0	0	0	2	1	0	0	0	0	0	0	0	0
265	0	1	3	2	6	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
267	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
268	0	1	0	1	2	0	0	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
274	0	1	0	2	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
275	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
278	1	1	0	1	1	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
322	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
329	1	1	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
332	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
335	1	1	3	2	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
342	0	2	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
343	0	2	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
344	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
369	3	1	3	6	6	0	1	1	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0	0
381	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
401	0	1	9	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
402	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

404	0	2	1	2	0	2	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
411	0	1	6	1	0	2	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
418	1	2	3	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
420	0	0	0	2	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
423	1	1	0	1	1	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
431	2	2	2	1	5	0	1	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0
434	0	1	3	1	0	2	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
436	0	0	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
456	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0
473	0	0	1	1	0	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
480	0	0	1	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
482	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
485	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
491	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
522	1	2	0	2	1	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
531	1	1	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
535	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
538	0	1	0	2	6	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
552	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
553	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
554	2	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
558	1	1	0	1	1	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
559	0	2	4	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
561	0	0	0	1	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
562	0	1	2	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
563	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
564	1	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
565	0	0	0	1	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
568	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
569	0	2	5	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
604	0	1	11	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
614	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
627	0	0	0	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
630	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
635	0	0	0	1	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
651	0	0	1	1	0	0	0	0	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
656	0	2	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
667	0	2	4	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
670	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
682	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
685	0	0	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
690	0	2	3	0	1	1	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
695	0	1	0	1	1	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
700	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
708	0	0	0	1	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

710	0	1	0	1	1	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
720	0	3	0	1	1	0	0	1	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
728	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
729	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
736	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
745	1	1	2	2	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
748	0	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
750	0	4	0	1	1	0	0	0	0	0	0	2	0	1	1	0	0	0	0	0	0	0	0	0	0
756	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
761	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
762	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
763	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
764	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
766	1	3	0	1	1	0	0	1	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0
772	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
773	0	2	5	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
774	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
777	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
782	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
787	0	1	1	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
796	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
797	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
803	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
805	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
809	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
810	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
811	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
901	0	1	0	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
916	0	1	10	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
918	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
920	0	1	3	2	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
930	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
935	0	0	1	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
954	1	1	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
983	0	1	8	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1203	0	0	0	1	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1212	0	1	0	1	2	0	0	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1217	0	2	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
1220	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1221	0	2	12	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1238	0	2	2	2	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
1240	0	1	0	2	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

1252	1	0	1	1	1	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1257	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1277	0	0	1	1	0	0	0	0	0	0	1	1	0	2	0	0	1	0	0	0	0	0	0	0	0
1280	2	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1293	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1294	0	3	0	1	1	0	0	1	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
1295	2	2	11	6	7	1	1	2	0	0	3	0	0	0	0	2	0	0	1	0	0	0	0	0	1
1298	1	1	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1304	4	2	5	4	2	0	3	1	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0
1310	3	2	4	3	10	0	2	1	0	0	5	0	0	1	0	2	1	0	0	0	0	1	0	0	1
1312	1	1	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1319	2	2	3	4	4	0	1	1	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	1
1334	0	0	0	1	1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1455	0	2	1	1	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
112D2	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
123D2	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16D2	0	2	2	2	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
198D2	1	3	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
280D1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
318D1	0	2	2	1	0	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
336D2	1	3	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
413D1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
444D2	0	2	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
472D3	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
483D1	0	0	1	1	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
483D2	1	0	0	1	2	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
487D2Y	0	2	1	1	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
724D2	0	2	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
68D1	1	1	0	2	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0

Supplementary Table 3l **Overview of the genome, metabolome and inter-bacterial inhibition studies**

Bacterial isolates are displayed by their identification number in first column. First row and from left to right, BGCs; total count of biosynthetic gene clusters. node; sum of retrieved nodes from metabolites analysis. U. nodes; sum of unique nodes to one isolate. Ave. Anta. size; average in the sizes of halos of inhibitions for a producer strain. Nb. Anta.; number of inhibited strains. Ave. Sens. size; average in the sizes of halos of inhibitions for a sensitive strain. Nb. Sens.; number of times the strain is sensitive.

	BGCs	nodes	U. nodes	Ave. Anta. size	Nb. Anta.	Ave. Sens. size	Nb. Sens.
4	3	89	0	0,503	4	1,169	3
9	11	66	0	1,176	7	0	0
11	17	64	2	2,713	15	0	0
22	3	58	5	0	0	0,629	1
29	8	97	7	6,937	19	0,648	2
31	10	90	1	0	0	0	0
50	9	30	1	5,641	14	0,301	2
53	0	102	0	0,383	2	3,677	18
60	3	40	3	0	0	4,065	16
61	3	67	1	1,115	3	0,69	12
63	40	41	2	18,271	46	0,415	1
65	4	52	7	4,681	16	0	0
68	7	121	0	12,794	22	0	0
70	2	126	0	0	0	0	0
71	7	74	0	6,175	20	0,65	6
73	4	134	0	0,124	2	0,591	1
74	6	50	0	3,777	8	0	0
76	8	105	1	0	0	0	0
77	7	43	2	1,605	7	0	0
79	4	49	1	0,07	1	0,136	1
81	3	84	2	5,4	11	0,086	1
83	14	125	0	1,233	4	0	0
85	3	109	3	0	0	7,609	17
96	8	53	5	0,176	1	1,256	8
100	5	120	0	0,62	2	0	0
101	1	64	2	0	0	3,325	14
102	7	161	1	0,358	1	0,597	1
105	2	139	2	3,668	11	2,364	6
107	36	155	10	0,152	1	0	0
122	4	97	1	3,039	10	0,609	2

127	8	139	2	6,936	14	0,493	6
131	16	71	3	0	0	0	0
133	6	96	3	0,997	3	3,588	9
135	16	183	5	0,376	1	0	0
136	30	111	1	1,166	2	0,61	1
137	4	48	1	2,077	7	3,494	13
140	2	120	0	0,053	1	2,4	6
142	9	82	3	0	0	0	0
147	9	107	0	0,913	7	0	0
149	8	150	5	0	0	0	0
151	2	109	0	0,28	2	0,37	4
154	2	156	0	14,471	36	0	0
157	1	138	3	3,944	13	0	0
166	4	130	4	0	0	1,466	16
170	4	66	0	3,21	17	0	0
172	5	67	0	0,62	2	0,348	1
179	3	99	2	2,466	8	2,414	9
180	3	102	3	0	0	0,036	1
181	3	44	0	8,275	17	6,043	11
186	9	145	1	0,795	4	0,127	1
187	30	192	1	1,236	3	0,079	1
189	19	35	0	0,877	7	6,469	16
190	4	70	1	0,555	2	2,218	7
214	2	49	7	0	0	0,09	1
217	4	32	0	0,878	2	0	0
219	4	60	0	0	0	0	0
224	2	89	1	0	0	1,303	6
227	3	113	2	1,086	3	0,646	8
231	8	119	0	0	0	0,338	1
236	1	49	0	0	0	10,08	31
239	7	66	1	1,018	6	0	0
240	4	107	1	1,19	3	0,298	5
241	1	86	0	1,002	3	0	0
258	8	77	2	3,657	10	0	0
264	39	93	2	0	0	1,256	2
265	19	131	0	0	0	1,368	4
267	4	69	1	1,119	2	0	0
268	7	116	1	0	0	0,473	1

274	6	143	0	0	0	0	0
275	4	72	2	0,686	3	0	0
278	7	77	1	0,408	1	0	0
322	2	97	0	0,787	8	0,517	2
329	6	71	0	6,624	14	0	0
332	2	97	11	2,58	2	6,101	25
335	10	73	68	2,226	8	0	0
342	5	109	1	0	0	3,039	10
343	5	154	0	0,289	2	3,157	6
344	2	87	2	1,568	3	7,843	20
369	26	74	7	2,69	12	0	0
381	4	146	0	0,247	2	0,306	1
401	17	112	32	30,802	50	0	0
402	3	142	0	0,532	2	0,111	2
404	9	155	2	0,305	1	0	0
411	12	108	6	0,587	2	0	0
418	8	93	1	0	0	0	0
420	5	149	1	1,232	5	0,05	1
423	7	53	0	1,186	4	0	0
431	18	56	2	0	0	0	0
434	9	73	1	0,761	2	0,101	1
436	3	68	8	0,721	4	0	0
456	4	180	2	0	0	6,264	8
473	5	54	4	0,863	3	0,113	1
480	4	53	1	0,178	2	2,433	18
482	2	149	0	0	0	0	0
485	3	80	2	1,081	2	21,234	38
491	2	173	1	0,284	3	0	0
522	9	124	0	0	0	9,276	14
531	7	93	0	6,598	30	0	0
535	4	122	8	0	0	8,334	16
538	14	139	0	0,964	4	1,817	5
552	1	145	1	0	0	0	0
553	3	83	1	1,217	10	1,855	10
554	4	149	0	0	0	0,233	1
558	8	191	3	0	0	0	0
559	8	31	0	1,53	4	0,276	4
561	4	47	1	0	0	11,94	28

562	7	128	0	9,045	18	0	0
563	1	111	4	0	0	7,194	27
564	5	162	0	2,483	6	0	0
565	4	145	3	9,906	19	0	0
568	2	156	0	0	0	0	0
569	11	84	13	9,143	26	0	0
604	15	77	0	0,941	5	0,086	1
614	3	66	0	0,07	1	2,722	9
627	3	77	1	1,858	5	5,552	13
630	2	114	3	0	0	0	0
635	4	151	15	3,541	11	0	0
651	6	133	1	0,314	4	0	0
656	5	121	3	0	0	0	0
667	8	72	7	0,131	1	0,159	1
670	2	132	1	0,376	2	2,441	8
682	3	105	6	18,553	34	0,361	2
685	3	122	1	1,167	5	0	0
690	11	69	16	4,499	34	0	0
695	7	167	0	0	0	0,422	1
700	3	96	0	0,387	1	0,428	1
708	5	122	3	0	0	0,226	1
710	5	140	3	0,197	2	0	0
720	9	76	1	0	0	0	0
728	4	95	1	1,417	4	7,143	10
729	1	102	1	0,02	1	15,683	28
736	6	221	7	0,075	1	2,45	11
745	8	82	1	2,603	10	0,07	1
748	4	56	0	0,349	2	8,499	19
750	11	98	2	0,289	1	0,614	2
756	2	73	1	0	0	20,039	34
761	2	150	0	0	0	0,69	7
762	4	74	1	0,627	2	1,666	6
763	2	46	0	0	0	0,775	10
764	2	87	1	0	0	0	0
766	10	122	2	0	0	0,358	1
772	1	43	3	0	0	9,042	20
773	11	72	24	1,472	3	0	0
774	3	85	0	0	0	11,067	12

777	5	86	0	0	0	10,991	20
782	3	121	1	0	0	6,23	11
787	5	121	4	0	0	0	0
796	2	84	1	0	0	1,408	6
797	1	73	1	2,788	7	0,606	6
803	1	85	0	0	0	11,442	18
805	3	164	1	0	0	10,044	11
809	2	82	0	0	0	4,004	20
810	3	56	3	1,276	4	1,53	7
811	3	68	4	2,72	7	2,816	8
901	5	115	8	2,538	8	0	0
916	13	75	2	0,923	5	0,161	2
918	3	108	0	0,313	4	0	0
920	9	163	14	5,295	16	0	0
930	4	89	4	6,7	17	3,938	18
935	5	54	1	0,258	3	0	0
954	6	113	1	0	0	0	0
983	13	27	1	0,08	1	0,396	4
1203	5	129	0	0	0	0	0
1212	7	47	2	0	0	0,08	1
1217	5	128	2	0	0	0	0
1220	3	159	1	0	0	0,871	1
1221	19	129	0	0	0	0,469	2
1238	8	143	1	1,255	4	0	0
1240	6	72	4	2,052	5	1,278	5
1252	8	103	3	0	0	0	0
1257	2	76	0	0	0	1,755	7
1277	7	131	0	0	0	0	0
1280	4	130	0	0,951	3	0,211	2
1293	3	106	2	0	0	5,269	41
1294	9	59	4	6,019	13	0	0
1295	40	54	0	1,133	5	0,985	6
1298	6	46	0	0,209	1	0,038	1
1304	25	62	16	0,853	3	0	0
1310	39	63	3	0,707	4	0	0
1312	6	132	0	0,318	2	0,107	1
1319	22	159	17	4,823	14	0,342	2
1334	4	134	0	2,68	10	6,383	20

1455	7	114	2	0,138	2	5,295	19
112D2	3	107	4	4,11	11	0,254	2
123D2	3	126	0	3,267	10	0	0
16D2	8	129	4	0,154	1	0	0
198D2	7	76	0	0,362	1	0,109	2
280D1	3	123	0	0	0	2,141	15
318D1	8	122	0	0	0	4,614	25
336D2	7	61	0	2,837	6	0	0
413D1	2	131	0	7,004	21	0	0
444D2	5	67	7	0	0	2,204	5
472D3	3	72	3	3,08	11	0,186	2
483D1	6	285	20	2,982	11	1,175	3
483D2	6	70	2	2,196	6	0,953	2
487D2Y	7	134	1	2,395	9	0,372	1
724D2	5	107	0	1,132	3	3,491	5
768D1	7	150	1	0	0	0,106	1

Abbreviations

ABBA	antagonistic bacteria-bacteria interactions assay
A. thaliana	*Arabidopsis thaliana*
BGCs	biosynthetic gene clusters
$Comp^{(-)}$	*in vivo* depletion of the 13 highly competitive stains
hserlactone	homoserine lactone
nrps	non-ribosomal peptides
n.s.	not significant
PCoA	principal coordinate analysis
PUFA	Polyunsaturated fatty acid
$Sens^{(-)}$	*in vivo* depletion of the 13 highly sensitive stains
t1pks	type I polypeptides
t2pks	type II polypeptides
t3pks	type III polypeptides
transpks	trans-AT polypeptides
vs.	*versus*

Acknowledgement

I would like to thank my Ph. D. mentors, Dr. Stéphane Hacquard and Prof. Dr. Paul Schulze-Lefert, for offering me the opportunity to conduct my Ph. D. project. I appreciate all of their contribution of time, advice and funding. Being mentored in tandem was a rich, dense and fruitful learning experience for me. I am very thankful for both that have contributed equally and complementary to my intellectual and scientific awareness and particularly thankful for our ternary interaction to mature my Ph.D. project.

I would like to thank my TAC committee members, Dr. Kenichi Tsuda and Prof. Dr. Alga Zuccaro for their time, advices and all the constructive feedbacks. I was very humbled interacting with both during our TAC meetings.

I would like also to thank all of Dr. Max Crüsemann and Dr. Till Schäberle for conducting the metabolomic study and Dr. Thorsten Wille and Dr. Axel Hamprecht for conducting the screen against multi-drug resistant bacteria.

I am thankful to Prof. Dr. Sheng-Yang He and Prof. Dr. James Cole and Prof. Dr. James Tiedje for hosting me in Michigan State University. I particularly thank James Kremer for hosting my stay in MSU, for all of his scientific contribution and for the great shared experience in East Lansing. I am particularly also thankful to Benli Chai for his help and dedicated time to the development of the bioinformatic pipeline that I used in this study.

I would like to thank my Ph. D. defense committee members, Dr. Sebastian Fraune, Dr. Eric Kemen, and Prof. Dr. Gunther Döhlemann for their time, interest and constructive comments.

I would like to thank all of Ruben, Barbara and Thorsten for their countless bioinformatic tips, Stijn for his precious wet-bench advices and Yang for providing the bacterial strains in a very early stage. I would like to thank Kathrin and Jack for investing time in reviewing my Ph. D. document and correcting my typos. I would like also to thank all of Brigitte, Anna Lisa, Shiji, Paloma, Simone, Amelia, Masayoshi, Thomas, Ka-Wai, Rafal and Henning for the stimulating group meetings. I thank all the people from the department for providing a delightful and stimulating atmosphere during Tuesday seminars.

Last but not least, I would like to thank both of my families that are across the Mediterranean sea, Amel, Bettina, Djamel and Walter for their encouragement and involvement for my well-being during these years. I am very grateful to my wife, Anina, for her faithful support and encouragement to pursuit my goals and for all the support and patience during the late stage of my Ph. D.

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit − einschließlich Tabellen, Karten und Abbildungen − , die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie − abgesehen von unten angegebenen Teilpublikationen − noch nicht veröffentlicht worden ist, sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde.

Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Paul Schulze-Lefert betreut worden.

Köln, der 22.05.2017