UNIVERSITÀ DI PARMA

UNIVERSITA' DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN

"Biologia evoluzionistica ed Ecologia"

CICLO XXXVI

From plant microbiota to nitrogen-fixing symbiotic bacteria: clues for novel solutions in sustainable agriculture

Coordinatore:
Chiar.mo Prof. Pierluigi Viaroli

Tutor:
Chiar.mo Prof. Alessio Mengoni

Dottorando: Lisa Cangioli

Anni accademici 2020/2021 – 2023/2024

PhD dissertation

From plant microbiota to nitrogen-fixing symbiotic bacteria: clues for novel solutions in sustainable agriculture

Lisa Cangioli

Supervisor: Professor Alessio Mengoni

16 May 2024

Lisa Cangioli
From plant microbiota
to nitrogen-fixing symbiotic bacteria:
clues for novel solutions in
sustainable agriculture
PhD dissertation, May 2024
Supervisor: Prof. Alessio Mengoni

University of Florence
Department of Biology
Via Madonna del Piano, 6
Sesto Fiorentino, 50019, Florence

Preface

This thesis is the result of a three-years "journey" into the plant microbiome, from the complexities of the microbial communities to the intimate relationships between strains and plant genotypes. The central focus has been to understand the connections (through microbial ecology, ecological genetics and molecular approaches) between the microbiome and the host, to rationally develop novel approaches for low input agriculture.

Many actors were playing a role in this journey; the microbiome change, depending on the host genotype and the environmental conditions, even a single microbial species can show dramatic differences among strains in the behaviour with the plant. Which are the players? Microbial taxa, microbial genes, plant genotypes, edaphic conditions, climate, agronomic practices, and even randomness.

The chapters of this thesis are aimed at presenting the investigations and obtained results done part of such players.

- *Chapter 1* is dealing with microbial taxa and plant genotypes asking the question about the specificities of the plant microbiomes with respect to plant species (and varieties) and the way to improve its services to plant in terms of bioinoculant use.
- *Chapter 2* is moving to a specific taxon, where genotypes of single strains and plant genotypes display interactions and single bacterial genes in epistatic combinations contribute to plant phenotypes.
- *Chapter 3* enters into the genetic details of such combinations and investigates the ecological genetics of genetic transfer of symbiotic and nitrogen-fixation genes between plant symbiotic bacteria, which bring to strains with different symbiotic properties.
- *Chapter 4* continues the journey into the molecular aspects of the genetics below the differences in symbiotic behaviour of bacterial strains, presenting the results of investigations on epigenomic signatures of mutualistic interactions and strain-by-strain differences.

The appendix is then reporting two case studies where methods and expertise gained during this "journey" have been applied to phytoremediation.

Finally, the works in progress and the drafts of future publications are reported.

Riassunto

La mia tesi di dottorato è il risultato di un lungo viaggio durato tre anni alla scoperta del microbioma delle piante, partendo da un'analisi generale delle complessità presenti all'interno delle comunità microbiche del suolo fino alla dissezione delle relazioni più intime che si instaurano tra ceppi batterici e piante ospiti con genotipi diversi. Il focus principale è stato cercare di comprendere le connessioni presenti tra microbioma e pianta ospite al fine di sviluppare nuovi approcci eco-sostenibili da utilizzare in agricoltura attraverso l'ecologia microbica e aspetti genetici e molecolari.

In questo viaggio molti attori hanno preso parte e svolto ruoli fondamentali; il microbioma della pianta è principalmente influenzato dal genotipo della pianta ospite e dalle condizioni ambientali, ma anche l'alterazione di una singola specie microbica può determinare conseguenze drastiche che si ripercuotono nel comportamento di alcuni ceppi batterici nei confronti della pianta. Ci siamo chiesti quindi chi sono i protagonisti principali; i taxa microbici? O forse i geni microbici? Oppure potrebbero essere i genotipi delle piante, le condizioni edafiche, quelle climatiche, le pratiche agronomiche applicate al suolo, o magari la casualità?

Nel cercare di chiarire tutti questi punti interrogativi che ci siamo posti, il mio viaggio alla scoperta del microbioma delle piante è iniziato planando sugli aspetti più generali per poi atterrare nei sottili e delicati meccanismi che stanno alla base delle nostre domande. Infatti, la mia tesi di dottorato può essere suddivisa in quattro principali capitoli che mirano a presentare i principali protagonisti delle interazioni della pianta con il suo microbioma.

Nel primo capitolo, abbiamo indagato sulla specificità dei microbiomi in relazione a diverse varietà di piante della stessa specie. In particolare, la nostra attenzione è rivolta al genotipo della pianta e alla composizione tassonomica del relativo microbioma. Questo filone della tesi ha come scopo applicativo finale quello di sviluppare nuove strategie di "*plant-microbiome engineering*", ovvero di ingegnerizzazione del microbioma delle piante, con il fine di creare comunità microbiche sintetiche da testare in vitro e in campo che siano in grado di aumentare la produttività della pianta e favorirne la crescita anche in condizioni ambientali avverse o non ottimali. Lo sviluppo di nuovi bioinoculanti è divenuto ormai imprescindibile per mitigare gli effetti dei processi di domesticazione delle piante e gli impatti dei cambiamenti climatici. Infine, l'aumento della produttività delle attuali colture agricole è necessario per far fronte a una popolazione mondiale in continua crescita.

Il secondo capitolo, invece, si concentra su un particolare gruppo funzionale di batteri che si trova nel suolo e che possiede particolari abilità che lo rendono utile come strumento biotecnologico nelle pratiche di agricoltura sostenibile, i rizobi. Particolari specie di rizobi, infatti, sono in grado di instaurare relazioni di simbiosi con piante leguminose e di fissare l'azoto atmosferico in forme direttamente assimilabili dalle piante come nitriti, nitrati e ioni ammonio grazie all'attività dell'enzima nitrogenasi. Il processo di fissazione biologica dell'azoto da parte dei rizobi avviene in particolari strutture definite "noduli radicali" che si formano a livello dell'apparato radicale delle piante. Questa relazione di simbiosi tra pianta ospite e ceppo batterico si instaura in seguito a un dialogo molecolare che avviene tra i due partecipanti, il quale è talmente specifico da permettere associazioni esclusive tra specifici genotipi della pianta ospite e specifici genotipi batterici. In questo capitolo, infatti, l'obiettivo è stato quello di dissezionare le relazioni di simbiosi tra pianta e rizobio in modo tale da comprendere quali sono le forze che guidano questo processo altamente specifico. Il modello pianta ospite-ceppo batterico studiato in questa tesi è costituito da *Sinorhizobium meliloti* e da piante di

erba medica (*Medicago sativa* L.), in quanto piante foraggere di interesse agricolo utilizzate per la produzione di mangimi animali.

Il capitolo terzo entra in merito dei dettagli genetici di tali combinazioni pianta-rizobio e indaga sul trasferimento genetico di geni responsabili del processo di simbiosi e di fissazione dell'azoto che avviene in alcuni casi tra ceppi diversi della stessa specie di batteri simbionti. Le informazioni genetiche per questi processi sono contenute sul megaplasmide pSymA di *Sinorizobium meliloti*, il quale possiede un genoma tripartito costituito dallo stesso plasmide, dal cromide pSymB e da un cromosoma. La mobilizzazione del megaplasmide pSymA è resa possibile grazie al processo di coniugazione che talvolta avviene in modo del tutto naturale in presenza di particolari condizioni. Quali siano le condizioni che permettono l'evento di coniugazione è un capitolo ancora tutto da scoprire. È stato reso noto che a monte della formazione del pilus coniugativo sia presente una cascata regolatoria costituita dal sistema *rct* che, in condizioni standard, si trova in stato di repressione ad opera di uno specifico inibitore chiamato RctA. I meccanismi molecolari che portano all'attivazione di questa cascata e, conseguentemente, al processo di coniugazione batterica sono tutt'ora ignoti. Con lo scopo di comprendere le condizioni necessarie affinché avvenga lo scambio di materiale genetico, numerose prove preliminari sono state allestite, ed è stato messo a punto un protocollo atto a rilevare l'attivazione di geni selezionati appartenenti al sistema *rct* tramite la creazione di un sistema reporter basato sulla *Green Fluorescen Protein* (GFP).

Nel quarto e ultimo capitolo della tesi si indaga sugli aspetti molecolari della genetica che si celano dietro alle differenze riscontrate nel comportamento simbiotico dei ceppi batterici, andando quindi a studiare i meccanismi epigenetici che regolano le interazioni mutualistiche tra pianta ospite e ceppo batterico e che determinano differenze fenotipiche tra i singoli ceppi batterici. Lo studio della metilazione del DNA di diversi ceppi di *Sinorhizobium meliloti* è stato reso possibile tramite l'analisi bioinformatica dei profili di metilazione ottenuti dal sequenziamento su piattaforma Pacific Bioscience (*Single Molecule Real-Time, SMRT, Sequencing*).

Nell' appendice finale è riportato un ulteriore caso in cui lo studio dei microbiomi è stato applicato al processo di *phytoremediation*.

Infine, sono presenti i draft dei lavori svolti e ancora in fase di stesura per la pubblicazione.

Contents

Chapter 1 3

Plant – microbiome interactions 3

1.1 What is plant microbiome? 5

1.2 Importance of plant-microbiome interactions in agriculture systems 7

1.3 Impact of global change and effects of domestication on the plant microbiome 7

1.4 Plant- Microbiome engineering as biotechnological tool for sustainable agriculture 9

1.5 Microbiome engineering using synthetic microbial communities 11

Chapter 2 14

Symbiosis 14

2.1 From modern agriculture to sustainable agriculture: a compulsory shift 15

2.2 Biological Nitrogen Fixation and its economic and ecological advantages 16

2.3 Symbiotic Nitrogen Fixation 18

2.3.1 Rhizobia and legumes: the partnership 19

2.4 *Sinorhizobium meliloti* 20

2.5 *Sinorhizobium meliloti – Medicago sativa* as a model plant interaction 22

2.6 Specificity of rhizobia-plant interaction: molecular aspects of the nodulation process 25

2.7 Genes involved in nodule formation and nitrogen fixation process 26

2.8 Not only symbionts: comparative genomics and phenotyping of the bacterial genera *Ensifer* spp. associated with the plant 27

Chapter 3 29

Movement of the symbiotic genes between rhizobia. A glimpse from S. meliloti pSymA transfer 29

3.1 Horizontal gene transfer and conjugation process in rhizobia 29

3.2 Regulation mechanisms of the symbiotic megaplasmids conjugation process 31

3.2.1 RctA-mediated repression system in *Sinorhizobium meliloti* 32

3.3 The question and the aims of our study 34

Chapter 4 35

How epigenetics affects rhizobial strains differentiation and symbiosis? 35

4.1 DNA methylation in bacteria 35

4.2 Unravelling DNA methylation role in genomes 36

Perspectives 38

Bibliography 40

Appendix 47

Appendix – Phytoremediation: an interesting ability of plant microbiome to adapt to harsh environmental conditions 47

Work in progress 53

I. Synergistic effect of rhizobia and plant-growth promoting nonsymbiotic bacteria on alfalfa growth

II. Study of the regulatory cascade involved in the megaplasmids conjugation in the model plant symbiotic bacterium *Sinorhizobium meliloti* by the creation of a GFP-based reporter system

III. The impact of synthetic microbial consortia, fertilization regimes and maize growth stages on taxonomic and functional microbiome under temperate climate conditions

IV. Genomic characterisation of bacterial strains composing an effective synthetic community applied as biofertilizer for maize, wheat and tomato

Overall activities 50

Publications 51

Conferences 54

Courses 56

Chapter 1

Plant – microbiome interactions

Outline

In the last decades, the importance of plant-microbiome interactions has become very clear and many studies have been focused on the rhizosphere due to the importance of the soil habitat of plants. The rhizospheric microbiome of plants plays an important role in plant growth and development, plant health and plant productivity[1]. Plant-microbiome can also mitigate the effects of domestication and intensive agriculture systems that led plants to be less resilient to environmental changes because they have impacted plant microbiomes for several reasons.

With the aim to exploit the potential benefits of plant microbiome, researchers have addressed their studies to unravel the composition of these microbial communities and to dissect the interactions that rule their relationship with plants. In order to understand soil microbiome and associated alterations in the molecular contents on community level induced by ecological interactions, the combinatorial multi-omics data resulted from metagenomic, metatranscriptomic, metaproteomic, and metabolomic have been fully investigated. The targeted sequencing of bacterial 16S rDNA, eukaryotic 18S rDNA, fungal ITS (Internal Transcribed Spacer), as well as whole-metagenome shotgun (WMS) sequencing help us to identify the variety of microbiomes and can be used to investigate how genomic information is converted into structures and functions in the interactions between plants and their microbial partners[2]. Thanks to the revolution in omics, gene-editing and sequencing tool, many biotechnological and ecological alternative to standard practices in agriculture have been proposed but many complexities and challenges in plant-microbiome engineering still remain[2]. In fact, despite technologies have demonstrated that plant responses to environmental changes are influenced by the interactions with its microbiome, little is known about the structure of microbial communities and how the environmental factors affect microbial networks.

In this context, plant-microbiome represents an innovative biotechnological tool to be applied directly to agricultural systems to meet the necessity to improve the production in a more sustainable and eco-friendly way and to face the global changes impacts. Microbiome engineering provides different strategies to be applied directly in field. Nowadays, two main approaches exist to engineer the plant microbiome: direct inoculation of exogenous beneficial microorganisms and re-inoculation of ex-situ enriched indigenous beneficial microorganisms. Principally, the application of microbial bioinoculants as biofertilizers or stress protection products and the creation of synthetic microbial communities (SynComs), or the microbial seeds coating, represent innovative strategies for a future more sustainable agriculture[1,3–5]. Concerning the first approach, the most difficult step of testing bioinoculants is moving from laboratory-scale to field application[6]. In particular, our knowledge concerning non-model plant is very poor and needs to be implemented with new studies concerning the emerging crops and landrace that represent important genetic resources. In this Chapter, we reported two our reviews, one that explains the main limitations related to model plants study and the economic and ecological relevance of

prioritizing the study of plant microbiome in non-model plants, and the other one that explains some tools that we analysed in the second group of the papers we reported in this chapter that refers to the creation and application of SynComs. These microbial communities can be considered an innovative solution to manipulate the soil and plant microbiomes directly in natural field conditions for practical purposes[7], but one of the most challenging aspects of their application is their maintenance in soil and their long-term stability under adverse conditions and under competition with the native soil microbiome.

Under the light of these evidence, in this Chapter we focused our attention on two different approaches and papers were grouped for the following aspects: i) the analysis of the response of two staple crops rhizospheric microbiomes in relation to different variables and ii) the evaluation of the effects of applying synthetic microbial communities, directly in field conditions.

i) In the first group of works, we analysed the microbiome of two important crops for the global food request, wheat and potato, with the aim to understand if environmental and intrinsic factors could affect the composition of the plant microbial community associated to the roots. In fact, in this Chapter of the thesis we reported two papers in which we studied the taxonomical composition of plant microbiome in response to different plant genotype and input of fertilization comparing old and modern wheat varieties, and in different phenological stages in relation to pedoclimatic conditions for the same cultivar of potato. We performed a 16S rRNA gene amplicon sequencing analysis of rhizospheric soil collected from wheat and potato, and we observed that: a) concerning the wheat, a variety-specific response of rhizosphere bacterial community was detected, both in relation to taxa (Nitrospira) and metabolic functions involved in the nitrogen cycle, suggesting that variety can influence the cycling of plant-assimilable nitrogen, b) concerning the potato, the rhizomicrobiota resulted affected by both the site of cultivation and the plant stages, with the presence of site-specific taxa and functional profiling of the microbiota, which could be further exploited as biomarkers of origin and traceability of the products.

ii) In the second group of works, we tested a collection of plant-growth-promoting bacteria isolated from a non-model crop (African rice) and applied to a model crop (Asian rice) with positive effects on several aspects. We also developed and tested different SynComs under field conditions taking part to some activities of the Horizon 2020 research project SIMBA (Sustainable Innovation of Microbiome Applications in Food System) for maize and tomato. In the end, we also developed and tested another SynCom contributing to the Italian Ministry of Agriculture research project Micro4Legumes, where a consortium of nitrogen-fixing symbiotic rhizobia and auxin producing bacteria strongly increased alfalfa production. The overall results obtained from the application in field of SynComs demonstrated that these microbial communities are able to improve plant resistance to biotic and abiotic stresses and to mitigate the negative effects of domestication and climate global change.

1.1 What is plant microbiome?

Scientists have been long fascinated by the intricate relationship between plants and microorganisms, and this topic continues to be an attractive research field[8]. In fact, plants are host to a rich and diverse community of microorganisms, including bacteria, fungi, archaea, and viruses, collectively known as *plant microbiome*. Components of the plant microbiome can establish different types of association with the plant, spanning from symbiotic (mutualistic), to commensal and even pathogenic.

These microorganisms inhabit every accessible plant tissue, from the roots to the leaves and reproductive organs, and engage in a multifaceted dialogue with their host[9–11]. Recently, plant microbiome can be seen as an extension of the plant genome[1,12], and the collective genomes of the host and its microbiome represent the holobiome[7]. Consequently, host plant and its associated microbiota have been defined as a "*holobiont*", in which evolutionary processes (e.g. natural selection) between plants and microorganisms takes part to the stability of the system[7,11,13]. Plant microbiome composition is dynamic, and it is influenced by different factors over the course of a plant's lifetime, such as plant developmental stage, plant and microorganism genotypes, and growth conditions. These circumstances can bring to shifts in plant microbiome composition (Figure 1). Generally, the *core microbiome* of a species consists of microbial community that associate as a result of evolutionary selection, whose ubiquitous members can carry essential genes for host fitness and are persistent in almost all the individuals of the same plant species [11,14]. In contrast, the accessory microbiome is shaped by variable biotic interactions between the microbial community and its host; the accessory microbiome can help plants adapt to specific environmental condition and then can play relevant roles in resistance and resilience to environmental perturbation[11,12].

Indeed, plant microbiome influences different plant processes, such as growth and development, nutrient acquisition, disease resistance, and environmental responses through dynamic and complex interactions. Indeed, plant microbiome is a key determinant against biotic and abiotic stresses[10,12,15].

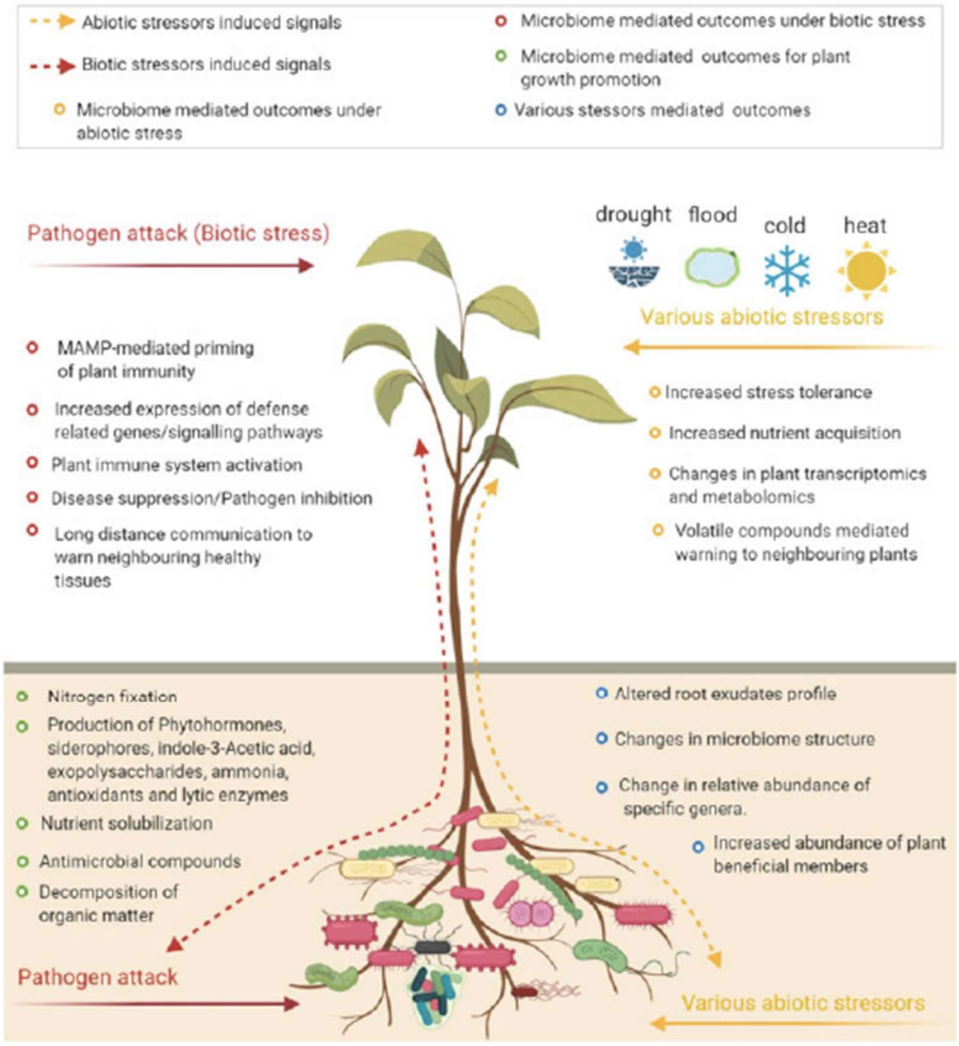


Figure 1. *Modulation of various plant functions by plant-associated microbiome. The plant and its associated microbiome are usually in a mutually beneficial association where the plant provides common goods for the microbial community, and in return the associated microbiota promote plant health and growth. A variety of biotic and abiotic stresses can elicit changes in the plant-associated microbiome. These changes in the microbiome in turn benefit the plant via a variety of mechanisms enabling the host plant to respond better to those stresses*[6].

Plants and soil microorganisms extensively interact with reciprocal impacts on fitness. Plants need nutrients present in the soil to grow but, sometimes, these nutrients are available but inaccessible for the plants. On the other hand, microbes constantly need shelter from harsh environmental conditions and nutrients supplied by plants. In particular, the rhizospheric microbiome, the microbial biomass that lives in the region around plant roots, defined as "rhizosphere" more than a century ago by Lorenz Hiltner[16], is directly impacted by the presence of the plant root and its secretions, and it is largely dependent on the abiotic environmental factors (such as pH, the temperature, the soil texture, the moisture content, the nutrients availability)[7,9,12]. Similarly, agronomical practices such as the utilization of chemical fertilizers and pesticides are able to model and alter the composition of microbial biomass present in the soil[12]. Finally, even small difference in plant genotype can influence the recruitment of soil microbial communities, leading to differences in microbial composition[17].

1.2 Importance of plant-microbiome interactions in agriculture systems

In the last decades, the availability of sufficient food is one of the most incoming problems due to the ever-growing population and the limited food-production resources. Human activities, such as deforestation and degradation of soils because of intensive agricultural practices and climate changes, have led to a decreased area of arable soil; this situation does not fit with the mandatory target to increase crop productivity from 60 to 100% by the year 2050 to meet the request of food for the estimated future global population (9.7 billion). To meet this goal, farmers apply chemical fertilizers that are not healthy for the soil and suitable for the food chain[18].

Plant microbiota is sculpted by the intimate interactions between the host plant and its associated microbes. Dissection of plant-microbiome interactions is highly complicated, but it is necessary to better understand their related functions. Deciphering these interactions and analysing the plant microbial communities could enhance our understanding on plant microbiome as modulator of host with its abiotic environment, improving plant productivity and crop yield[7]. In effect, considering the holobiont, the microbial genes present in the microbiome augment the host's own repertoire and act to improve the host's adaptation to environmental perturbations or, in some cases, prevent it from doing so[1,10].

Focusing our attention on beneficial microbes, they can also help plants to mitigate the effects of the most significant factors that affect agricultural productivity, as for instance salinity and drought[11]. In these stressful conditions, plant-growth-promoting microorganisms (PGPMs), such as plant-growth-promoting bacteria (PGPBs) and plant-growth-promoting fungi (PGPFs), enhance plant resistance adopting various strategies such as phytohormone production, lowering ethylene oxide levels, upregulation of dehydration response, and the induction of genes encoding antioxidant genes. A relevant role is played by the microbiome of the plant rhizosphere, where beneficial bacteria are termed as plant-growth-promoting rhizobacteria (PGPRs)[18–20].

Hence, the exploitation of PGPMs can represent a promising strategy to be considered as complement in agriculture systems to increase yield and resistance of crops.

Indeed, PGPMs are now being tested, and in some cases already used, in some modern agriculture practices to produce food crops and enhance crop productivity with low input from agrochemicals. Recent studies have shown how the application of bioinoculants composed by complex consortia of beneficial microorganisms, the so-called *synthetic microbial communities* or, simply, *SynComs*, can lead to an improved crops yield[21–25]. For instance, a very hopeful alternative to nitrogen input from chemical fertilizers is the inoculation of SynComs composed by nitrogen-fixing bacteria. In this regard, rhizobia are model species since they fix atmospheric nitrogen by establishing a symbiotic relationship with plant roots.

1.3 Impact of global change and effects of domestication on the plant microbiome

Plant domestication refers to the co-evolutionary process based on the implementation of breeding programs in which humans select and modify wild plant species based on plant traits useful for adaptation to cultivation and human-derived purposes. Through the domestication process, wild species have been subject to both artificial and natural selection which are resulted in the independent and repeated evolution of similar plant phenotypes known a *domestication syndrome*[26].

One of the most significant goals of plant domestication process is to increase the crop yield to provide humans with a continuous and constant food supply. In effect, domestication processes have selected the most productive genotypes and the most suitable crop lines. On the other hand, domesticated plants are characterized by less ability to interact with or adapt to the surrounding environment, and they required much more measures that their wild parents, such as fertilization, irrigation, and other protective treatments against phytopathogens.

As an unintended consequence of these alterations in plant microbiome and due to the fact that the selected crop cultivars might have lost some of the genetic traits needed to recruit host-specific microbiota, selected crop lines are less resilient to biotic and abiotic stresses as compared with their wild relatives[26,27]. In conclusion, domestication and intensive agriculture systems have impacted plant microbiomes for several reasons, including (i) reduction of genetic diversity and genomic modification in cultivated plants, (ii) modification of key traits controlling microbiome assembly and structure and (iii) alteration of agricultural habitats and potential microbial source communities[26].

Environmental conditions have been altered also by climate changes. The atmospheric concentration of the carbon dioxide, and other greenhouse gasses, is still increasing and the increase of global temperature has been associated with it. Because of higher temperatures, soil water content in some areas is decreasing, leading to drought and salinity conditions. Climate change has also affected the root exudates composition with an increased carbon concentration to the root zone. The different C/N ratio affects the microbial communities leading to an alteration in species composition and, consequently, in the related functions. The impacts of climate change have globally several repercussions on plants and their associated microorganisms. Under suboptimum conditions, plant physiological processes can be altered or diminished, and they affect plant growth and development. While warming and drought have negative effects on plant performance and crop production, on the other hand, they also affect the microorganisms that live in the soil associated with their hosts. In fact, climate change induces alteration in composition and activities of plant microbiome, leading to consequence to host function[18].

Climate variability has already reduced the crop productivity world-wide, and its negative effects are expected to intensify in coming decades. Therefore, new approaches to mitigate the negative effects of climate change in plants are urgently required[28,29].

Recent studies showed how the plant-associated microbiome can modulate and alleviate the impacts of biotic and abiotic stresses on plant fitness; in the short terms (years to decades), is mainly the plant microbiome to drive the adaptation of plant to climate change, whereas in the long term (century to millennia), the adaptation of plants will be driven by the co-evolutionary interactions between plant microbiome and its host[28].

A deeper knowledge of plant and its microbiome response to climate change and a better understanding of the strategies and the molecular mechanisms that plant microbiome adopts to mitigate the negative effects offers an important perspective to predict the climate change impacts on plant primary productivity and aids in developing management and policy tools to improve the resilience of plant systems. To harness the mitigation processes of plants and its associated microbiome to abiotic stress, it is essential to understand the molecular machinery, as well as metabolic pathways and regulatory genes. To understand stress mitigation strategies, the identification of stress-responsive multigenic traits and exploration of linked markers for stress related genes are the current focus[18,28].

1.4 Plant- Microbiome engineering as biotechnological tool for sustainable agriculture

One of the biggest challenges that modern agriculture needs to deal with is accelerating agricultural productivity to satisfy the increasing food request due to a constant growing human population exploiting the available resources (land, germ plasma, water) without compromising the terrestrial, aquatic, and air ecosystem[3]. Microbiomes represent a potential biotechnological tool to be applied to agricultural systems to meet the need to produce more food in the most sustainable and environment-friendly way. The achieving of this goal without damaging the agricultural system is challenging. Innovations that involve the use of microbiome, such as the development of microbial bioinoculants as biofertilizers biocontrol or stress protection products, the creation of synthetic microbial communities, or the microbial seeds coating, may represent the future of sustainable agriculture[1,3–5].

Nowadays, it is known that every plant interaction with the surrounding environment is mediated by interactions with the microbiota. Despite the advances in sequencing technology over the past decade have showed that plant responses to the environment, including abiotic and biotic stress, are partially mediated by interactions with its microbiome, our understanding of how microbial communities are structured and how the environmental and intrinsic factors modulate microbial network remains a crucial question mark. Plant microbiome research has highlighted a complex network of genetic, biochemical, physical and metabolic interactions among the plant, the associated microbial community and the surrounding environment[3]. Those interactions play a key role in shaping the assembly of microbial plant-associated community and its taxonomic and functional structure in plants, and, combined with the adoption of specific agricultural practices and certain environmental factors, they positively influence and modulate the beneficial traits of plant microbiome, such as plant growth and crop productivity, plant health, stress tolerance and nutrients acquisition, which can lead to higher crop yields[30]. For this reason, plant-microbiome interactions represent a promising research field to explore for plant breeding and crop improvement[3,11,26,31].

To investigate the complex taxonomy structure of microbial plant-associated communities, the field of plant microbiome research has mostly relied on the amplicon-based sequencing, such as 16S rRNA gene sequencing for bacteria and ribosomal Internal Transcribed Spacer (ITS) sequencing for fungi. Although the taxonomic composition of plant microbiome is now almost clear, it remains to be explored and understood other microbiome features, such as activity and functional capacity, through the inclusion of meta-transcriptomics and shotgun metagenomics, and the link between plant microbiome data and plant physiology, genetics, metabolism, and other host processes[10], such as signalling events that mediate microbe-host interactions, along with the inclusive quantifiable systems biology approaches[3].

In the last decades, many studies have encouraged plant microbiome engineering to provide a set of highly advanced technologies and strategies to apply directly in field. Principally, two different approaches exist to engineer the plant microbiome: direct inoculation of exogenous beneficial microorganisms and re-inoculation of ex-situ enriched indigenous beneficial microorganisms.

In particular, the root microbiome can be engineered by soil conditioning (to improve the microbiome diversity, for example, using legumes to enrich native diazotrophic microorganisms) or substrates / non-microbial biostimulants (including new organic soil amendments and root exudates to attract the beneficial microbiota). Alternatively, inoculating microbial strains into a preexisting community to modify the structure of microbial communities could also be used to achieve such goals[32,33],

including rhizosphere microbiome (bacterial competitiveness engineering), synthetic microbiome (genetically engineered microbes inoculation to host plant), in situ microbiome manipulation of the native microbial community in their native context, and plant mycobiome (the optimization and improvement of beneficial plant-fungal interactions)[34–36]. However, these new strategies have to face up all the potential biotic and abiotic perturbations that can influence or hinder their role[3,37].

One of the most difficult step of studying beneficial effects of bioinoculants is moving from laboratory-scale to field application[6]. Concerning model plants microbiome, our knowledge is becoming clear, whereas little is known about non model plants, which includes most of the emerging crops and landraces which hold important genetic resources. In this review, we explain limitations related to model plants study and the economic and ecological relevance of prioritizing the study of plant microbiome in non-model plants.

Under these premises, in this thesis we analysed the native microbiome of two staple crops, wheat and potato, aiming to understand the i) relationships between plant genotype, input of fertilization and the rhizosphere microbiome, comparing modern and old varieties of wheat, and ii) the successional stages of the rhizosphere microbiome in relation to pedoclimatic conditions. Published papers on these topics are reported below:

i) Cangioli, L., Mancini, M., Napoli, M., Fagorzi, C., Orlandini, S., Vaccaro, F., & Mengoni, A. (2022). Differential Response of Wheat Rhizosphere Bacterial Community to Plant Variety and Fertilization. International Journal of Molecular Sciences, 23(7), 3616. doi: https://doi.org/10.3390/ijms23073616

ii) Cangioli, L., Mancini, M., Baldi, A., Fagorzi, C., Orlandini, S., Vaccaro, F., Mengoni, A. (2022). Effect of Site and Phenological Status on the Potato Bacterial Rhizomicrobiota. Microorganisms, 10 (9), art. no. 1743. doi: https://doi.org/10.3390/microorganisms10091743

Results from these analyses indicate that old and modern wheat varieties, having differential use of nitrogen input, also modulate the rhizosphere microbiome in relation to the process of nitrogen cycling and defined possible microbiological variables related to pedoclimatic conditions, which could be further exploited as indicators of origin and traceability.

Moreover, to discuss the need for translating knowledge on model crops microbiome to emerging crops a focus paper was published, where research priorities and methodological proposals are presented (iii):

iii) Vaccaro F., Cangioli L., Mengoni A., Fagorzi C. (2022). Synthetic plant microbiota challenges in non-model species. Trends in Microbiology 30 (10), pp. 922-924. doi: https://doi.org/10.1016/j.tim.2022.06.006

Some of the ideas and tools presented in this focus paper have then been tested and discussed in the following papers (iv, v, vi) as presented in the next section (1.5).

1.5 Microbiome engineering using synthetic microbial communities

Plant-growth promoting bacteria (PGPBs) research field has been harnessed since years and aims to the identification of beneficial bacterial strains, which can confer plant resistance, alleviate the negative effects of biotic and abiotic stressors, and increase the crop productivity[14,26]. Nowadays, there are numerous cases of single PGPB strain applied on field as biocontrol agent to increase plant resistance to phytopathogens, including wheat[38] and rice[39]. However, it becomes essential to deeper investigate the rhizosphere microbial ecology and plant microbiome composition, and study the effects of exogenous PGPBs inoculation in the soil and their interaction with native soil microbial community and the host[14].

First of all, it is necessary to select highly competitive organisms with good colonisation abilities[5].

PGPBs are isolated from soils and used as inoculum; currently, different microbial inoculants are on the market. Unfortunately, the effectiveness of a single PGPB strain application depends, principally, on whether this exogenous microorganism can survive in the competition with the native microbial communities[11], on climate, on soil type and on environmental factors[28]. The competitiveness of the bacterial inoculum can be enhanced by inoculating crops with consortia of bacteria, or Synthetic microbial communities (SynComs), instead of single strain. Starting from single microbial cultures, previously selected and tested for their antagonistic activity toward pathogens and plant-growth promoting traits and for excluding they can reciprocal inhibit, consortia are obtained mixing the strains collection. The required steps in constructing a SynCom include the selection of microbes members, culturing the core microbes, the optimization of microbial interactions, and assessing the efficacy of the SynCom[40]. This advanced microbiome engineering strategy increases both the survival and the activity of the inoculated microbiota and also mitigates the interactions of consortia with indigenous microorganisms, with the plant host, with the microclimate (temperature and humidity), and soil chemical and physical structures (pH, water, nutrient availability). Nevertheless, only a small fraction of soil microbial taxa are culturable, the most natural microbiomes remain uncultivated[41]. A solid alternative to modify the soil native plants microbiota is represented by the introduction of genetically modified microorganisms that have acquired the resistance genes to adverse conditions, but the introduction of these modified microorganisms into the open environment is subjected to stringent national biosafety regulations[32,42]. Recently, also the seed coating has been shown to promote microbial colonization and thus could improve the efficiency of SynCom compared to soil application, due to priority effects.

However, bacterial synthetic communities (SynComs) represent a powerful and promising approach to dissect pant-microbiome interactions reducing the complexity of natural systems, and excluding confounding environmental effects, in a controlled setting. In fact, one of the most advantageous aspects of SynComs is represented by the possibility to manipulate by the addition, elimination, or substitution of strains, to detect the main PGP trait of individual bacterial culture, such as plant growth promotion, diseases and stresses resistance in plant hosts (Figure 2). Moreover, SynComs enable not only to decipher the mechanism in plant-microbiome interactions *in vivo*, but also to verified the experimental output directly in natural field conditions for practical purposes[7].

For designing of synthetic communities, it is necessary an extensive study of plant microbiota and an elaborate culture collection[43].

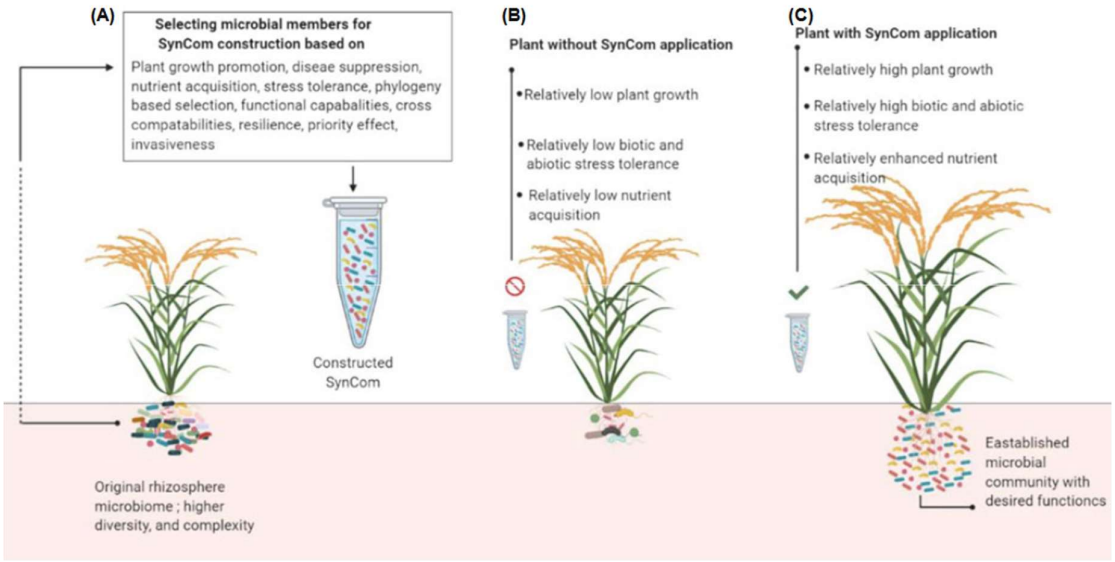


Figure 2. *Construction of synthetic microbial community (SynCom) with desired functions to enhance plant growth and tolerance to biotic and abiotic stressors. (A) The SynCom members are selected (in the above depiction from rhizosphere) based on various criteria such as growth promotion, tolerance to various kinds of biotic and abiotic stress, and nutrient acquisition. SynCom is defined and has a relatively simple community structure and low microbial diversity as compared to the original rhizosphere microbial community. (B) SynCom can be designed and potentially applied to target plants which are either nutrient-deprived or under various kinds of biotic or abiotic stress. (C) Applying the SynCom comprising of defined microbial members with desired functions could potentially affect plant growth and health by either promoting nutrient acquisition or providing tolerance against biotic and abiotic stress via various mechanisms*[7].

To be considered an effective microbial consortium, it should be more able to tolerate fluctuation in environmental conditions than a single strain and increase plant growth and resistance to biotic (i.e., suppression of phytopathogens) and abiotic stresses. However, one of the major challenges in the application and engineering of a SynCom still remains its maintenance in soil and its long-term stability under competition with the native community for limited niches and in adverse conditions. To develop stable and effective SynCom for agriculture, novel approaches and using different dimensions of omics data, such as genome and microbial profiling data, will be required[40].

We outlined the need for rationalize researches on plant-associated microbiomes in a review paper (iv) and while testing a collection of PGPB from a non-model crop (African rice) to a model crop (Asian rice) (v) we also developed (vi) and tested a SynCom under field conditions (vii, viii) under the activities of the Horizon 2020 research project SIMBA (Sustainable Innovation of Microbiome Applications in Food System, https://simbaproject.eu/) for maize and tomato. Moreover, we also developed and tested a SynCom under the Italian Ministry of Agriculture research project Micro4Legumes (https://www.sinab.it/ricerca/il-microbioma-vegetale-simbionte-come-strumento-il-miglioramento-delle-leguminose-foraggere), where a consortium of nitrogen-fixing symbiotic rhizobia and auxin producing bacteria strongly increased alfalfa production (ix).

iv) Fagorzi, C., Passeri, I., Cangioli, L., Vaccaro, F., Mengoni, A. (2023). When biodiversity preservation meets biotechnology: The challenge of developing synthetic microbiota for resilient sustainable crop production. Journal of Sustainable Agriculture and Environment, 2 (1), pp. 5-15. doi: https://doi.org/10.1002/sae2.12038

v) Bianco, C., Andreozzi, A., Romano, S., Fagorzi, C., Cangioli, L., Prieto, P., ... & Defez, R. (2021). Endophytes from African Rice (Oryza glaberrima L.) Efficiently Colonize Asian Rice (Oryza sativa L.) Stimulating the Activity of Its Antioxidant Enzymes and Increasing the Content of Nitrogen, Carbon, and Chlorophyll. Microorganisms, 9(8), 1714 doi: https://doi.org/10.3390/microorganisms9081714

vi) Gullì M., Cangioli L., Frusciante S., Graziano S., Caldara M., Fiore A., Klonowski A.M., Maestri E., Brunori A., Mengoni A., Pihlanto A., Diretto G., Marmiroli N. and Bevivino A. (under submission). Combined metagenomic and metabomic analyses to unlock the potential of co-applied SynComs and novel amendments from residual biomasses for a more sustainable agriculture.

Chapter 2

Symbiosis

Outline

Since the invention of the Haber-Bosh synthesis process to produce ammonia, the production of chemical nitrogen fertilizers improved and allowed the global population to increase exponentially. On the other hand, the industrially produced nitrogen (N) fertilizers require non-renewable energy resources for their production, and they represent a risk for the human and the environmental health. An emerging alternative to reduce the use of chemical products is represented by the biological nitrogen fixation (BNF); this process is performed by a large variety of different organisms, such as aquatic organisms, free-living soil bacteria or bacteria associated with plants. In particular, some soil bacteria, named rhizobia, are able to establish a symbiosis relationship with leguminous plants and, thanks to the nitrogenase enzyme, they can fix the atmospheric nitrogen (N_2) into ammonia (NH_3), which can be directly assimilated by plants. Rhizobial nitrogen fixation provides 40-48 million tonnes of N into agricultural systems each year and its relevance in sustainable agriculture promoted the production of rhizobia inoculants. The application of bioinoculants as biofertilizers meets the need to increase the crops yield in a more sustainable way. Moreover, rhizobia contribute to plant health and development with other plant-growth promoting (PGP) traits, such as resistance to biotic and abiotic stresses and defence from phytopathogens attacks.

One of the most studied rhizobia-leguminous models is represented by *Sinorhizobium meliloti*-*Medicago sativa* (alfalfa). *S. meliloti* is an α-proteobacterium with a multipartite genome comprising three main replicons: the chromosome and two megaplasmids, pSymA and pSymB. pSymA megaplasmid carries most of the genes involved in nitrogen fixation process and related to the symbiosis instauration. *M. sativa* is a perennial leguminous plant cultivated as forage crop in many countries of the world. The symbiosis relationship occurs in specific structures, named nodules, formed at root level. The mechanism which allows the symbiotic interaction between rhizobia and alfalfa is a multi-step process in which many biomolecules are involved to establish a molecular crosstalk between bacteria and plants. This process is highly regulated by a synchronism of factors, such as the specificity of plant and bacteria compounds released in the rhizosphere and the presence or absence of inducers in the surrounding environment. In fact, different strains of *S. meliloti* show in their genome differences in relation to the symbiotic genes, both in terms of presence and regulation, and these differences are reflected in terms of quality of symbiosis and host preference. Furthermore, recent studies demonstrated that the efficiency of the symbiosis and the number of formed nodules are directly determined by plant genotype x rhizobial genotype interactions, and also influenced by the interactions with the environment.

In this chapter, we reported two reviews in which we promote research on genotype x genotype interaction between rhizobia and host plant in order to design more specific synthetic symbiotic nitrogen-fixing microbial communities to be applied in sustainable agriculture, and we discuss about the main actors which take part in the establishment of symbiosis. Moreover, to evaluate whether

plant and rhizobial genotypes influence the initial transcriptional response of rhizobium following perception of a host plant, RNA sequencing study of three *S. meliloti* strains elicited by three different alfalfa varieties root exudates as host plants and luteolin was performed. Analysing the transcriptome, we observed that the response to root exudates involved hundreds of changes, of which 35% were influenced by the strain genotype, 16% were influenced by the plant genotype, and 29% were influenced by strain-by-host plant genotype interactions. Most of the upregulated genes were conserved in all three strains, suggesting that the *S. meliloti* lineage was adapted to rhizosphere colonization before gaining the genes required for symbiotic nitrogen fixation. In conclusion, these results showed that the initial transcriptional response of rhizobium to legumes is influenced by both partners genotypes and their interaction, providing novel genetic insights into the genetic basis of rhizobium-plant interactions for improvement of legume growth in agricultural ecosystems.

Another question mark that we wanted to discuss in this chapter was the presence within the same taxonomic group of symbiont and nonsymbiont strains. Many studies investigated how symbiotic rhizobia evolved, but all of these forced the evolution in laboratory conditions and not along strains phylogeny. The current rhizobia classification is based on phenotypic and genetic features, in particular takes into account the 16S rRNA gene phylogeny. The bacterial genus *Ensifer* contains some important nitrogen-fixing bacteria, as well as nonsymbiont species. Through a comparative genomic study, we observed that two different clades were highlighted, the “symbiotic” clade and the “nonsymbiotic” one, and several phenotypic differences between clades were found too. These results supported the hypothesis to divide these clades into two main subgroups and that genetic features aside from the classical symbiotic N_2-fixation genes contribute to the evolution of symbiotic potential.

2.1 From modern agriculture to sustainable agriculture: a compulsory shift

Since the agriculture has been practised, the practices were organic without using synthetic fertilizers or other agrochemicals. This era of the *ancien régime agriculture* ended with the introduction of the Haber-Bosch synthesis process to produce ammonia combining its gaseous constituents, nitrogen and hydrogen, under high temperature and very high pressure. This process was rapidly developed into industrial scale production to synthetize the commercial nitrogen fertilizers needed to grow most of the world’s hybrid crops[44]. A century after its invention, the Haber-Bosch process is still the basis for the production of more than 100 million tons of artificial nitrogen fertilizers per year[45]. This process served as the "detonator of the population explosion", enabling the global population to increase from 1.6 billion in 1900 to 9.8 billion by 2050; billions of people would never have existed without it[46,47].

Current modern agricultural practices often supply in excess the limiting factors for plant growth and productivity, especially water, mineral nitrogen, and mineral phosphate, and actively control pests[44]. The three hallmarks of modern agriculture are the control (i) of crops and their genetics, (ii) of soil fertility via chemical fertilization and irrigation, and (iii) of pests (weeds, insects, and pathogens) via chemical pesticides, but, nowadays, it is essential to develop a new long-term and multidisciplinary research program with agricultural methods that can feed a growing world and still preserve the vital services provided to humanity by the world’s natural ecosystems[44]. To meet these goals and to sustain a sufficient food production to satisfy the ever-growing human population request, sustainable agriculture represents an environmental and healthy friendly solution. Sustainability is defined as

"the successful management of resources for agriculture to satisfy changing human needs while maintaining or enhancing the quality of the environment and conserving resources" (https://pdf.usaid.gov/pdf_docs/pnaaz324.pdf) and it can be described as the agriculture that is managed toward greater resource efficiency and conservation while maintaining an environment favourable for evolution of all species. One of the driving forces behind agricultural sustainability is the effective management of N in the environment. Industrially produced nitrogen (N) fertilizers deplete non-renewable energy resources and pose human and environmental hazards. Moreover, judicious management of N inputs into cropping systems is a prerequisite for sustainable agricultural practice. A very successful way for manipulation of N inputs that results in farming practices that are economically viable and environmentally prudent is the use of biologically fixed nitrogen. Symbiotic systems can be a major source of N in most cropping systems[48].

2.2 Biological Nitrogen Fixation and its economic and ecological advantages

The economic and environmental costs of the heavy use of chemical N fertilizers in agriculture are a global concern. Sustainability considerations mandate that alternatives to N fertilizers must be urgently sought[49]. A promising strategy to reduce the use of chemical fertilizer is represented by the biological nitrogen fixation (BNF), which is the process that allows the synthesis of organic nitrogenous compounds from atmospheric nitrogen by certain microorganisms; in particular, all known nitrogen-fixing organisms (diazotrophs) are prokaryotes. Prokaryotes able to perform BNF include aquatic organisms, such as cyanobacteria, free-living soil bacteria, such as *Azotobacter*, bacteria that form associative relationships with plants, such as *Azospirillum*, and most importantly, bacteria, such as rhizobia, that form symbioses with legumes[50,51].

Species of *Azotobacter*, *Bacillus*, *Clostridium*, and *Klebsiella* represent some examples of heterotrophic bacteria that live in the soil and fix significant levels of nitrogen without the direct interaction with other organisms (free-living nitrogen fixers). These organisms find their own source of energy typically by oxidizing organic molecules released by other organisms or from organic matter decomposition. There are some free-living organisms that have chemo-lithotrophic capabilities and can thereby utilize inorganic compounds as a source of energy[52]. In Figure 3, it is shown that the ability to fix nitrogen is widely distributed across the bacterial domain[53].

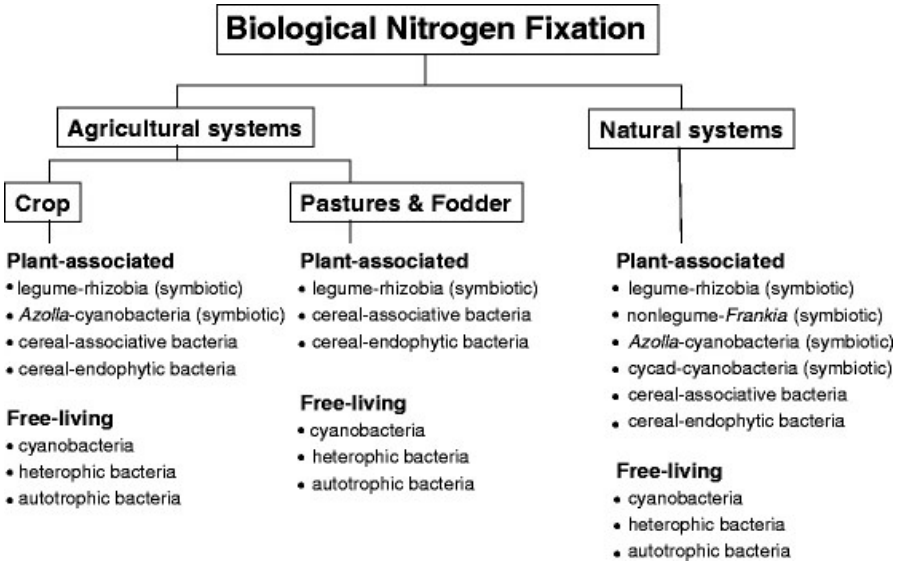


Figure 3. *Biological N_2-fixing agents in agricultural and terrestrial natural systems*[53].

Biological nitrogen fixation plays a key role in agricultural sustainability because nitrogen-fixing systems offer an economically and ecologically solution to reduce external N inputs and improve internal resources. Especially in symbiotic systems such as legumes and rhizobia, BNF represents the major N source, or a minor source supplementing inorganic N in non-symbiotic, nonlegumes such as cereals. In most agricultural settings, the primary source (80%) of biologically fixed N is through the legume nitrogen-fixing symbiotic bacteria from genera as *Rhizobium*, *Bradyrhizobium*, *Sinorhizobium* and *Azorhizobium*[48].

The nitrogen fixation process in these organisms is performed by the nitrogenase enzyme system, which, at 16 ATPs hydrolysed per N_2 fixed, carries out one of the most metabolically expensive processes in biology[54], catalysing the conversion of atmospheric nitrogen (N_2) to ammonia (NH_3). Plants can readily assimilate ammonia to produce the nitrogenous biomolecules.

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

A better knowledge of BNF on ecosystem and regional scales is essential to understand, predict and manage the consequences of multiple components of anthropogenic global change[55,56]. Since domestication of plants, nitrogen fixing species have played an integral role in cropping systems and have been prominently featured in rotations and intercropping, as alley crops, in pasture systems, as green manures, in agroforestry, and cover crops. Since the development of the Haber-Bosch process, implementation of legumes into cropping systems has declined due to the low cost of N fertilizer, high yielding cereal crops, and other economic and political forces. Today, this natural process has again a significant importance in world agriculture, but with a different aim: the demand for accurate determinations of global inputs of biologically fixed nitrogen is strong and will continue to be fuelled by the need to understand and effectively manage the global N cycle. The most important nitrogen fixing agents in agricultural systems are the symbiotic associations between crop and forage/fodder legumes and rhizobia, partnerships between plants and Frankia spp. or cyanobacteria, or from non-symbiotic associations between free-living diazotrophs and plant roots, during decomposition of crop residues, or as cyanobacteria crusts on soil. These N fixing systems can satisfy a large portion of their

own N requirements from atmospheric N_2, and that additional fixed N will be contributed to soil reserves for the benefit of other crops or forage species[57] (Figure 4). Annual inputs of fixed N are calculated to be 2.95 Tg for the pulses and 18.5 Tg for the oilseed legumes. Soybean (*Glycine max*) is the dominant crop legume, representing 50% of the global crop legume area and 68% of global production. Every year soybean is able to fix 16.4 Tg N, representing 77% of the N fixed by the crop legumes[53]. Finally, biological N2 fixation plays an important role also in desert lands because it is the major responsible for N input into desert ecosystems. The wild (naturally growing) leguminous plants living in arid regions are subject to severe environmental conditions.

In the end, it has been reported that wild legume-rhizobium associations are useful in providing a vegetational cover in degraded lands where desertification caused disturbance of plant-microbe symbioses, which are a critical ecological factor in helping further plant growth in degraded ecosystems. BNF is an efficient source of fixed N_2, which plays an important role in land remediation[58].

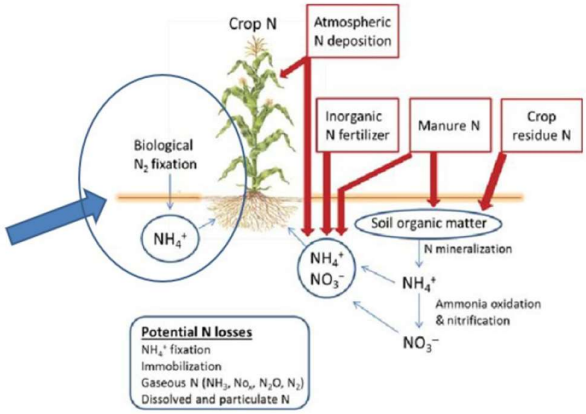


Figure 4. *The biological nitrogen fixation as N input for plant requirement. From: CAB Rev Persp (23):1-18.*

2.3 Symbiotic Nitrogen Fixation

The close association between microorganisms and plants, which can result in a symbiotic partnership, arises from the need for nutrients for microorganisms. The plant provides sugars from photosynthesis and in exchange for their carbon sources, microbes provide ammonia for its host. By far, the most important nitrogen-fixing symbiotic associations are the relationships between legumes and rhizobia. Alfalfa, beans, clover, cowpeas, lupines, peanut, soybean, and vetches represent the most relevant legumes used in agriculture, with soybean that tops the rankings of the most cultivated legume[59].

Rhizobium-legume symbioses provide more than half of the world's biologically fixed nitrogen, and it was reported that rhizobial nitrogen fixation introduces 40–48 million tonnes of nitrogen into agricultural systems each year[60,61].

In general, microorganisms have become popular models for addressing sociobiological questions related to exchange of nutrients and communication between different organisms[62]. The study of mutualistic interactions, microbial systems represent an optimal model, due to their readiness of manipulation, their short generation times, and the simple methods for tracking of resources[63].

The high economical relevance of the symbiotic interaction has promoted the production of commercial rhizobia inoculants (already widely used in agriculture), since this microbial interplay can provide a direct benefit for the host plant (Figure 5). Inoculation of leguminous crops with rhizobia

strains is known to considerably increase crop yield, mediated largely through symbiotic nitrogen fixation[62].

The effect of rhizobia with such PGP abilities, the so called plant growth promoting rhizobacteria (PGPRs)[64], is more evident under stressful conditions, since they allow to alleviate plant stress (for instance through the reduction of ethylene production) and increase nutrition (increasing root growth and fixing nitrogen)[65–68].

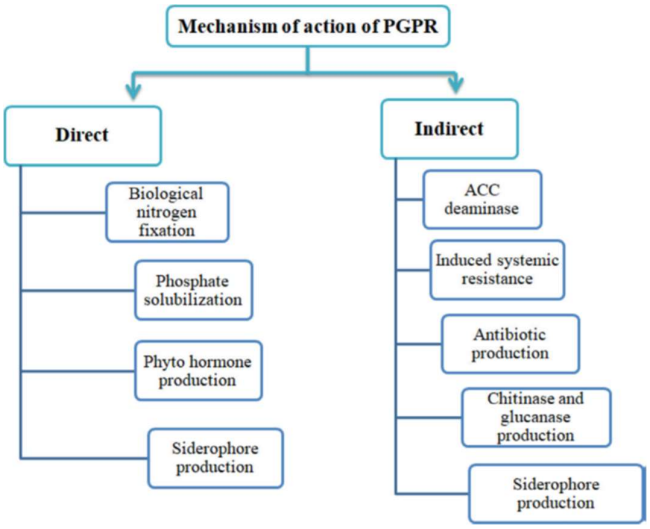


Figure 5. *Flow chart showing mechanism of action of PGPRs*[64].

Therefore, plant growth and agricultural yield is related not only to the plant genotype and the soil condition, but, especially for legumes, their associated and symbiotic microbiota also play important roles. As such, the selection of rhizobial strains resistant to harsh conditions, as for example water deficiency, and capable of alleviating metal phytotoxicity could be a crucial strategy to improve the yield of legumes in arid or in metal-contaminated soils.

2.3.1 Rhizobia and legumes: the partnership

In last decades, molecular researcher have been attracted by rhizobacteria, also classified as plant growth-promoting rhizobacteria (PGPRs)[69]. "Rhizobia" is a general term that indicates a large and diverse group of soil bacteria that are able to infect and colonize plant roots (*rhiza*, as root in Greek) and performing biological nitrogen fixation. When compatible leguminous plants are present, rhizobia may form symbiotic nitrogen fixing associations with them. Rhizobia is a polyphyletic group present in α- and β-Proteobacteria. While the β-rhizobia are mainly found in association with tropical legumes, the α-rhizobia appear to be more widespread and make symbiosis with tropical to temperate legumes including pasture, tree, and grain legumes. The α -rhizobia have received more research attention than the β-rhizobia; among the α-rhizobia, *Sinorhizobium* is likely the most studied genus, followed by the genera *Rhizobium, Mesorhizobium*, *Allorhizobium* and *Bradyrhizobium*[70]. Within Rhizobiaceae and the above-mentioned genera, other non-rhizobial species are present; for example, in the *Rhizobium* genus, there is *Rhizobium radiobacter*, commonly known as *Agrobacterium tumefaciens*[71,72]. The symbiotic nitrogen-fixing bacteria belonging to the Rhizobiaceae family are a genetically diverse and physiologically heterogeneous group interacting with the roots of

legumes to form nodules where atmospheric nitrogen is fixed. The genetic structure of rhizobial populations is associated with a high level of genetic diversity. In symbiotic rhizobacteria, the evolution of the population structure may be influenced by environmental conditions such as biological barriers to gene exchange, or geographical isolation, soil type and genotype of the host plant[54]. Presently, the classification of rhizobia consists of 98 species in 13 genera. The current classification is strongly influenced by polyphasic classification, considering phenotypic and genetic characteristics, with 16S rRNA gene phylogeny providing the backbone of the classification[73]. The advantage for rhizobia to make symbiosis with plants have extensively been discussed, as model in socio-microbiological studies and in relation to the evolution of conflicts and avoidance of cheating. In general, participating in symbiosis by founding a nodule can dramatically enhance the reproductive success of symbiont rhizobia. Even when some rhizobia inside some nodules lose the ability to reproduce, especially in nodules with indeterminate growth, it appears that every nodule contains large numbers of reproductively viable rhizobia[74]. Most of the genes required for nodulation are located either on large (500kbp to 1.5-Mbp) plasmids or on symbiosis islands. Nodules induced by rhizobia are of two general kinds, determinate and indeterminate. Indeterminate nodules are elongated and have a persistent meristem that continually gives rise to new nodule cells that are subsequently infected by rhizobia residing in the nodule. These newly infected cells, and the bacteria inside them, develop further and form new nodule tissue that actively fixes nitrogen.
This process results in a gradient of developmental stages, from the young meristem at the nodule tip to the older senescent tissue near the root. *Medicago sativa* (alfalfa), *Medicago truncatula*, *Pisum sativum* (pea), *Vicia* species (vetches), and *Trifolium* species (clovers) are used as models for studying the formation of indeterminate nodules. Determinate nodules are usually round and lack a persistent meristem and don't display a developmental gradient as do indeterminate nodules. Legumes that form determinate nodules are typically tropical in origin and include *Glycine max* (soybean), *Vicia faba* (bean), and *Lotus japonicus*[75].

2.4 *Sinorhizobium meliloti*

One of the most studied and best characterized rhizobia is *Sinorhizobium meliloti*. *Sinorhizobium meliloti* is an α-proteobacterium which belongs to the order of Rhizobiales[73], distributed and present in many temperate soil types, both in association with legumes or in a free-living form. *S. meliloti* is able to infect roots of *Medicago* species (*Medicago sativa* and its close relative *Medicago truncatula*) and develop a symbiosis. *S. meliloti* has been chosen by many international groups as laboratory model organism particularly interesting for its high genetic and phenotypic variability and for the multipartite nature of its genome. Consequently, the association between *Medicago sativa* (alfalfa), a model organism for plant nodulation, and *S. meliloti* is emerging as a worldwide leading model system for symbiosis and nitrogen fixation studies[76].

S. meliloti is used as model species for investigating plant-bacteria interactions, rhizobia-legume symbiosis and symbiotic nitrogen fixation. As for many other rhizobia, its genome is multipartite comprising three main replicons (Figure 6) with distinctive structural and functional features (Figure 7): in the model strain *S. meliloti* 1021, we have a 3.65-Mbp chromosome, and 1.35-Mbp pSymA and 1.68-Mbp pSymB megaplasmids, for a total length of 6.7-megabase (Mbp)[77]. In many other strains non-pSym plasmids of different sizes have also frequently been reported.

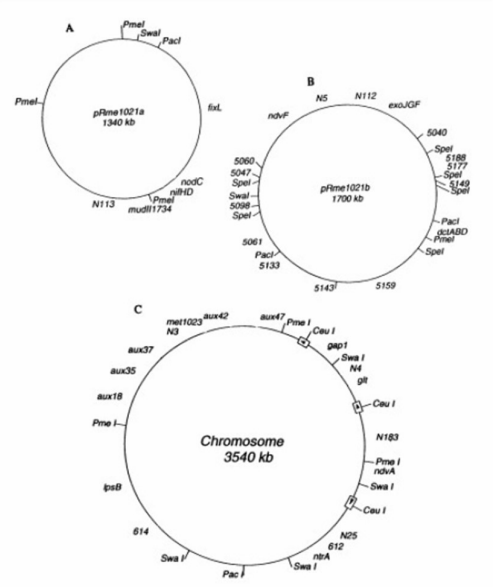


Figure 6. *Physical maps of the replicons of R. meliloti 1021. (A) pSym-a (10 3.7 kb); (B) pSym-b (1° 4.7 kb); (C) chromosome (1° 9.8 kb). Arrows indicate the orientations of the rrn loci. For all replicons, solid lines indicate the positions of target sites and/or transposon insertions anchored to the physical map; the positional order shown for other markers residing on the same restriction fragment is random*[78]*.*

The *S. meliloti* chromosome sequence carries, besides housekeeping genes, genetic information for mobility and chemotaxis processes, plant interaction (putative virulence genes), as well as stress responses. Functions could not be assigned for up to 41% of the *S. meliloti* chromosomal genes - 12% of these genes are unique, as they have no sequence analogy - because they were not found in other α-proteobacteria[76]. Contrary to expectations, the genome of *S. meliloti* is not highly reiterated. However, the *S. meliloti* genome contains a limited number of recently duplicated genes, but it contains many ancient duplications. This theory is confirmed because 42% of *S. meliloti* genes belong to 548 paralogous families, ranging from 2 to 134 genes per family; this high level of paralogy suggests that genome size has been little constrained during *S. meliloti* evolution, facilitating the acquisition of new adaptive functions for life in the soil and for symbiosis. Two megaplasmids that were acquired separately by an ancestor whose genome consisted of a single chromosome are pSymA and pSymB. pSymA seems to be the most recent replicon in the *S. meliloti* genome, as indicated by its distinctive GC% and codon usage, its paucity of Rhizobium-specific Intergenic Mosaic Elements (RIMES) and ABC elements. The pSymA megaplasmid mostly contains the genes related to the symbiosis interaction and to the nitrogen-fixation process and related to the symbiotic capabilities (as *nod*, *nif* and *fix* genes). These genes are liable to variations. In particular, the genes involved in microaerophilic growth during bacteroid development, those who determines the competitiveness inside nodules, their number for each plant, and those for the nitrogen metabolism can be in some strains completely absent or downregulated. This can of course determine a variation in the symbiotic phenotype of the bacterium, with an effect on the plant growth, giving a genetic base to the existence of cheater strains[77]. No essential gene could be predicted on pSymA; therefore, this replicon is clearly plasmidlike[76]. In *S. meliloti*, nodulation genes required for the synthesis and export of Nod factors are located on pSymA. Overall, sequence analysis suggests that the *S. meliloti* nod genes have two distinct origins: horizontal gene transfer (HGT), mediated by import of pSymA from an unknown bacterium, and resident gene duplication[79]. The chromosome of *S. meliloti* is that of a typical aerobic,

heterotrophic bacterium. Acquisition of pSymB chromid considerably extended the metabolic capabilities of the microbe by allowing it to metabolize a large variety of small compounds encountered in the soil or in the plant rhizosphere. An increased capacity in synthesizing surface polysaccharides, including exopolysaccharides (EPSs), lipopolysaccharides (LPSs), capsular polysaccharides (CPSs), and cyclic P-glucans, may also have significantly improved the colonization potential of these microbes. These compounds are encoded mainly by the chromosome and pSymB, and are crucial for successful plant infection, possibly by suppressing plant defence responses. The pSymB chromid is mainly harbouring genes related to transport and catabolism of carbon sources but it contains also some essential genes, including the arginine tRNA; then, pSymB has several chromosomal features[76]. Recently, the term "chromid" has been proposed to describe large replicons that contain not only plasmid-type replication systems but also essential genes for growth and survival[80].

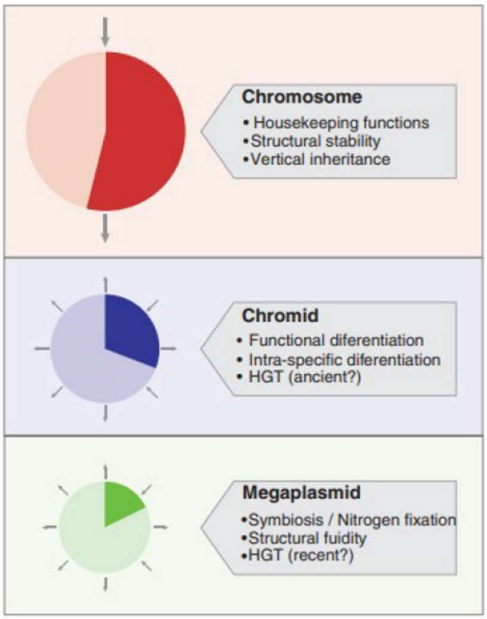


Figure 7. *Tasks and evolutionary differences of replicons in S. meliloti. Pie charts indicate the proportion of each replicon that is present in the DNA backbones. The arrows indicate the transmission mechanism: vertical inheritance (two arrows) or HGT (radial arrows)*[80].

2.5 *Sinorhizobium meliloti – Medicago sativa* as a model plant interaction

As reported in the previous paragraphs, symbiotic systems such as that of legumes and rhizobia can be a major source of N in most cropping systems. One of the most studied biological N fixation models is the *Sinorhizobium meliloti* strains and *Medicago sativa* symbiosis. *M. sativa* (commonly named alfalfa) is a perennial leguminous plant cultivated as forage crop in many countries of the world. It is native of temperate climate, but it is known to be tolerant to a large variety of temperature. It can be used as forage, for the high protein and vitamin content (important for the ruminants' nutrition), but also as biofuel and for soil quality improvement. Indeed, alfalfa cultivation naturally allows to enrich the soil with nitrogen after the depletion due to the previous crops (e.g. cereals) cultivation[77]. The mechanism which allows the symbiotic interaction between the nitrogen fixing rhizobia and leguminous plants is highly regulated. Indeed, the multi-step mutualist association consists in a

complex and specific signals exchange between both the organisms. In *S. meliloti*, the symbiotic process starts with the mutual recognition between the plant and the bacteria present in the rhizosphere, which is the soil area close to the roots exposed to a large variety of microorganisms, and so can represent an important source of nutrients for bacteria due to the presence of root exudates[77]. The symbiosis starts when flavonoids and other plant compounds (2-phenyl-1,4-benzopyrone derivatives, as for instance luteolin) - which can passively diffuse across the bacterial membrane - selectively bind a receptor protein in the bacterial cell (NodD), activating the expression of the bacterial nod genes, which allows the production and secretion of the Nod factor. The exact chemical structure of the Nod factor and the flavonoid secreted by plant, are the basis for host–symbiont specificity[77].

Other Nod factor stimulates cell divisions in the root, resulting in nodule organogenesis. The nod genes encode approximately 25 proteins required for the bacterial synthesis and export of Nod factor. Nod factor is a lipo-oligosaccharide signal consisting of a chitin backbone, four to five N-acetylglucosamine units in length, with a lipid attached to the nonreducing end and host-specific modifications on the backbone. When host plant perceives Nod factor, plant shows developmental changes in the early in the nodulation process, including root hair curling, membrane depolarization, intracellular calcium oscillations, and the initiation of cell division in the root cortex, which establishes a meristem and nodule primordium (Figure 8).

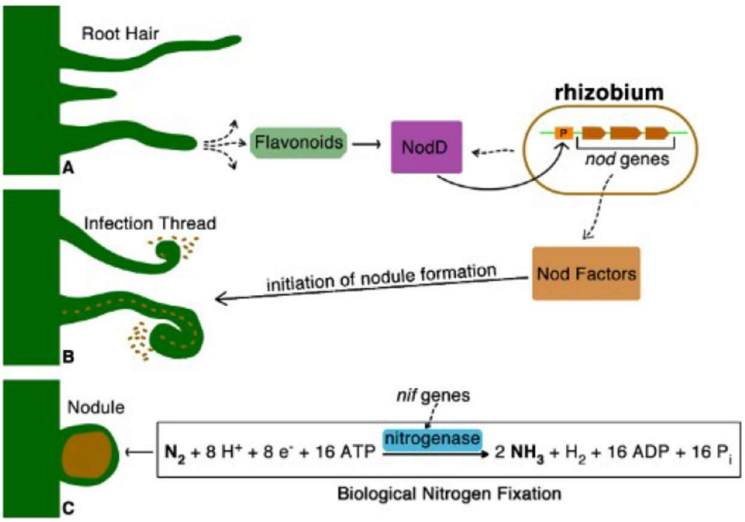


Figure 8. *Schematic overview of the nodulation process and biological nitrogen fixation*[81].

If the initial dialogue between plant and bacteria is successful, the root hair curls to trap a small number of bacteria. From this trap site, the root hair begins an inverse tip growth, forming a long and narrow passage, called the infection thread, in which the bacteria travel by continuously dividing at the leading edge. The infection thread traverses the entire length of the cell, eventually merging with the basal membrane to release the bacteria into the extracellular space. Additional infection threads then form in cells of underlying layers, repeating the process to deliver bacteria to their final destination, the root cortex (Figure 9). Other responses of rhizobia involve changes in the expression of nod genes and other genes important for rhizobia to compete effectively with other organisms for access to growth substrates emanating from the host root, to adhere to the root surface, and to become resistant to toxic substances such as phytoalexins secreted by the root. While the bacteria penetrate the root, host cells in the root cortex reacquire properties of stem cells, which form the

lateral organ nodule. When bacteria reach these nodule cells, they are internalized in an endocytosis-like process. The bacteria infect the nodule, enter the cytoplasm of plant cells and differentiate into a distinct cell type called a bacteroid, which can fix atmospheric nitrogen into ammonia. Alfalfa produces nodules that display an indeterminate developmental programme: a meristem at the tip of the growing nodule continually produces plant cells that are infected by bacteria, followed by differentiation of both the plant cells and the endosymbiont. As a result, all stages of development are present in a distal-to-proximal gradient within a single nodule. The meristem at the tip contains dividing plant cells that are not infected by bacteria and in older nodules there is a senescent zone in which both plant and bacterial cells degenerate[75,82–84].

Moreover, in a number of tropical legumes it seems well established that rhizobia do not infect via root hairs at all. In these plants, the bacteria appear to invade by penetrating the intercellular spaces between cells near the root surface, then move from the middle-lamellar region to the cytoplasmic membrane of dividing cortical cells. The nature of the penetration site for rhizobia and the functional role of root-hair deformation and bacterial entrapment, if any, are thus still uncertain[85].

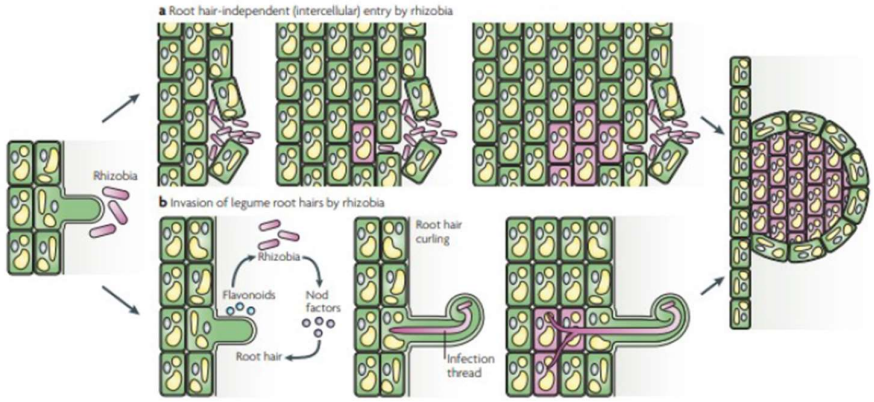


Figure 9. *Bacterial penetration of the root tissue can occur by two different strategies. In root hair-independent (intercellular) entry (a), disruption of the epidermal cell layer allows rhizobia to enter the root. This contrasts with the classic model of root hair invasion (b), where flavonoids released by legume roots trigger the synthesis of rhizobial Nod factors, which induce root hair curling, bacterial penetration at the centre of the infection pocket and division of cortical cells*[86].

2.6 Specificity of rhizobia-plant interaction: molecular aspects of the nodulation process

Rhizobia can bind plant root hairs through a two-step process. Direct attachment of bacteria to the surface of the root hair cell is mediated by a bacterial Ca^{2+}-binding protein, designated as rhicadhesin. The second step is the accumulation and anchoring of the bacteria to the surface of the root hair. This leads to so-called firm attachment.

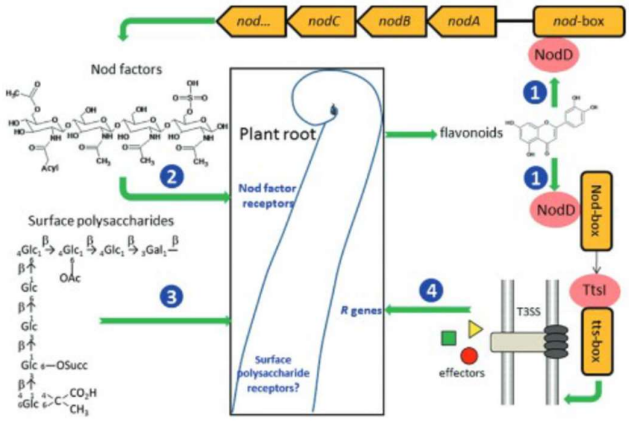


Figure 10. *Molecular determinants of host specificity during nitrogen-fixing symbiosis. (1) At the first stage, the plant produces flavonoid signals (such as the luteolin shown here, from M. truncatula) to free-living soil bacteria, activating the bacterial NodD proteins. NodD proteins bind to the conserved nod-box in the promoters of bacterial nodulation genes to induce their expression. (2) The nod genes code for enzymes for the synthesis of Nod factors. Secreted Nod factors are recognized on the plant surface by transmembrane Nod factor receptors in a strain- and ecotype-specific fashion. Modifications on the Nod factor such as the length and saturation of the acyl group determine host specificity. Activation of Nod factor receptors triggers growth changes in the root hair to trap a small number of bacteria, which would give rise to the entire population colonizing the resulting nodule. (3) Possibly downstream of Nod factors, rhizobia also use their surface polysaccharides (such as EPS from S. meliloti, depicted) to modulate host range. The plant receptor(s) are unknown but may resemble animal receptors for surface polysaccharides from bacterial pathogens. (4) In certain rhizobial strains, NodD also induces the expression of TtsI, which codes for a transcriptional regulator that binds to highly conserved promoter elements, called tts boxes, upstream of operons encoding the type III secretion machinery and effectors. Recognition of these effector proteins by R genes present in some ecotypes or varieties of plants limits host range*[83].

Nodulation process starts with chemotaxis of rhizobia towards one or more root exudate components. Once the bacteria are close to the root, they attach to the tips of root hairs and begins the transcription of nod genes mediated by flavonoid activated bacterial NodD proteins, which are members of the LysR family of transcription regulators. NodDs bind to conserved DNA motifs, known as nod boxes, found in the promoter regions of nod genes (Figure 10). NodDs from different rhizobial species respond to different sets of flavonoids. Since different legumes secrete different types of flavonoids, the ability to trigger NodD dependent induction of nod genes in response to a specific spectrum of flavonoids defines an early checkpoint for the legume-rhizobial symbiosis[83]. Thus, the evolution of host specificity involves both rhizobial and host genes[83,87].

Nod factors from different rhizobia share the same chitin-like N-acetyl glucosamine oligosaccharide backbone with a fatty acyl chain at the non-reducing end, but differ in their length of the backbone, the size and saturation of the fatty acyl chain, as well as additional modifications at either end, such as glycosylation and sulfation. Such decorations on the ends of the Nod factor play a crucial role in determining whether the Nod factor can be perceived by a specific host. Perception of the Nod factor signal in legumes is mediated by Nod factor receptors (NFRs), which are plasma membrane-localized

serine/threonine receptor kinases, containing LysM motifs in their extracellular domains. Corresponding to the Nod factor structure, NFRs are a host determinant of symbiosis specificity[83].

2.7 Genes involved in nodule formation and nitrogen fixation process

The Rhizobium genes essential for infection and nodule formation can be divided into two classes. The first group are genes that are not specific to symbiosis per se but are required for physiological adaptation to the environment, whether within the nodule or under free-living conditions. These include functions such as biosynthesis of amino acids and other molecules, carbon metabolism and respiration complexes for microaerobic growth, several sets of genes involved in the formation of the bacterial cell surface, such as genes determining the synthesis of exopolysaccharides (*exo* genes), lipopolysaccharides (*lps* genes), capsular polysaccharides or K antigens, and b-1,2-glucans (*ndv* genes). Mutations in these genes disturb the infection process to various degrees, such as the inability to elicit the formation of infection threads, resulting in the formation of non-fixing empty nodules. The second class consists of the nodulation (*nod* or *nol*) genes. Inactivation of the nodulation genes can result in various in planta phenotypes, such as the absence of nodulation, a delayed but effective nodulation, or changes in the host range[82].

- ***nod genes***

nod genes are responsible of starting the nodulation process and they are activated by specific flavonoids produced by plants. Some of the nod genes appear to be interchangeable for nodulation function between different species and biovars and are therefore designated as common *nod* genes. On the other hand, some nod genes are involved in the nodulation of a particular host and are hence called host-specific *nod* (*hsn*) genes. In most Rhizobium species studied to date, the nod genes reside on the large symbiotic plasmid pSymA, that also carry the *nif* and *fix* nitrogen-fixing genes[88]. When this sym megaplasmid was introduced into other Rhizobium species or into Agrobacterium tumefaciens, the transconjugants formed small ineffective nodules on Medicago sativa. This indicated that genes controlling all of the early steps of nodulation and host specificity are carried by the megaplasmid[88,89].

- ***nif genes***

nif genes are needed for the synthesis and activity of the nitrogen-reducing enzyme called nitrogenase. In *S. meliloti*, they map on a large plasmid called pSymA, along with other genes involved in nodulation (*nod*) and in nitrogen fixation (*fix*)[90].

- ***fix genes***

fix genes, like *nif* genes, are needed for symbiotic nitrogen fixation and induced in symbiosis, they are indispensable for the correct functioning of the nodule because they regulate the nitrogen-fixation process in the symbiosome[90].

- ***exo and lps genes***

A possible role of *exo* and *lps* genes in the determination of host specificity has been suggested, but no clear genetic evidence has yet been given that Rhizobium surface components are major determinants of host range specificity[88].

Different strains of *S. meliloti* display relevant differences in the genome and also in relation to the above-mentioned symbiotic genes, both in terms of presence (as part of the accessory genome fraction) and regulation[80] Such differences are also reflected in terms of quality of symbiosis and host preferences as early reported[91]. More recently, to demonstrate that interactions between rhizobia

and legumes are subjected to a context-dependent selection, Heath et al. showed how *S. meliloti* fitness is influenced by plant genotype x rhizobium genotype interactions. They studied the proportion of nodules inhabited by different *S. meliloti* strains in four plant generations demonstrating that rhizobia directly benefit from nodulation with plants, and these benefits are affected by plant genotype[92]. In another studies, still Heath et al. highlighted how the number of root nodules – a trait of the symbiosis – is directly determined by plant genotype x rhizobial genotype interactions, but it is also influenced by the interaction with environment. Indeed, the variation of N availability in the soil was shown to negatively affect the number of associations (nodules) between rhizobia and legumes in response to nitrate[93].

To investigate the genetic bases of strain-by-strain differences we performed a transcriptomic analysis of different *S. meliloti* strains under the treatment with root exudates and flavonoids from different alfalfa varieties. This study (vii), showed the presence of a large number of genes whose expression is modulated by the plant genotype and by also the specific bacterial genotype, highlighting the relevance of cross genome-wide interactions during plant-microbe symbiotic establishment. The take-away messages and proposals for using such information for developing ad hoc bioinoculants have then been presented in two review papers (viii, ix)

vii) Fagorzi, C., Bacci, G., Huang, R., Cangioli, L., Checcucci, A., Fini, M., ... & Mengoni, A. (2021). Nonadditive transcriptomic signatures of genotype-by-genotype interactions during the initiation of plant-rhizobium symbiosis. Msystems, 6(1), e00974-20. doi: https://doi.org/10.1128/msystems.00974-20

viii) Cangioli, L., Vaccaro, F., Fini, M., Mengoni, A., & Fagorzi, C. (2022). Scent of a Symbiont: The Personalized Genetic Relationships of Rhizobium—Plant Interaction. International Journal of Molecular Sciences, 23(6), 3358. doi: https://doi.org/10.3390/ijms23063358

ix) Cangioli, L., Checcucci, A., Mengoni, A., & Fagorzi, C. (2021). Legume tasters: symbiotic rhizobia host preference and smart inoculant formulations. *Biological Communications*, *66*(1), 47-54. doi: https://doi.org/10.21638/spbu03.2021.106

2.8 Not only symbionts: comparative genomics and phenotyping of the bacterial genera *Ensifer* spp. associated with the plant

How did symbiotic rhizobia evolve? Why within the same taxonomic group of rhizobia some are plant symbionts and others are nonsymbionts?

Several studies have shown that an effective symbiosis cannot be obtained by the transfer of symbiotic plasmids from rhizobia of the genera *Rhizobium* or *Sinorhizobium* to closely related non-rhizobia from the genera *Agrobacterium* or *Sinorhizobium*[94]. Other studies on beta-rhizobia have nicely deciphered the genetic requirements to transform a pathogenic species into a symbiotic one with experiments of symbiotic plasmid transfer followed by experimental evolution[76,95,96]. However, these experiments were done forcing evolution in laboratory conditions. Few information is on the contrary present on the evolutionary changes occurred along strains phylogeny[97].

The bacterial genus *Sinorhizobium* (syn. *Ensifer*) contains ecologically important nitrogen fixing symbionts of leguminous plants, as well as nonsymbiotic species. However, the evolutionary dynamics of symbiotic nitrogen fixation within this genus are unclear, and it remains an open question of whether the gain of classical symbiotic N_2-fixation genes is sufficient to allow a bacterium to fix

nitrogen. In general, the primary genes required for SNF (i.e., the *nod*, *nif*, and *fix* genes) are located within mobile genetic elements that include symbiotic islands and symbiotic (mega)-plasmids simplifying their spread through horizontal gene transfer[62,98,99].
In recent years, rhizobia classification has undergone major changes and revisions and several novel species have been described. The current classification of rhizobia, as for other bacterial groups, is strongly influenced by polyphasic classification, taking into account phenotypic characteristics as well as genetic characteristics, with 16S rRNA gene phylogeny providing the backbone of the classification. From 16S rRNA gene comparisons, it is apparent that the monotypic genus *Ensifer* is highly related to *Sinorhizobium* (Balkwill, 2005) and *Ensifer* as genus name was then used for all member of this taxon. Hence, the genus *Ensifer* provides an ideal model to further explore the differentiation of symbiotic bacteria from nonsymbionts. This genus comprises rhizobia such as *Sinorhizobim (syn. Ensifer) meliloti* and *Ensifer fredii*, as well as non-rhizobia like *Ensifer morelense* and *Ensifer adhaerens*, and many members have been extensively studied producing an abundant set of experimental and genomic data[100].

Under these premises, we performed a comparative genomic analysis, coupled with high throughput metabolic phenotyping (Phenotype Microarray) to investigate the monophyly of this genus and of symbiotic and nonsymbiotic strains. Results (paper x) indicated that symbiotic *Ensifer* strains are genomically distinct from the nonsymbiotic ones and that within the symbiotic clade the megaplasmid carrying symbiotic genes possibly evolved several times.

From this analysis, a taxonomic follow up was then performed in 2022, by Kuzmanovic et al., and the unification of the genera *Ensifer* and *Sinorhizobium* in Opinion 84 of the Judicial Commission was no longer justified by current genomic and phenotypic data. Therefore, *Ensifer* and *Sinorhizobium* genus are now considered as separate genera[101].

x) Fagorzi, C., Ilie, A., Decorosi, F., Cangioli, L., Viti, C., Mengoni, A., & diCenzo, G. C. (2020). Symbiotic and Nonsymbiotic Members of the Genus Ensifer (syn. Sinorhizobium) Are Separated into Two Clades Based on Comparative Genomics and High-Throughput Phenotyping. Genome biology and evolution, 12(12), 2521-2534. doi: https://doi.org/10.1093/gbe/evaa221

Chapter 3

Movement of the symbiotic genes between rhizobia. A glimpse from S. meliloti pSymA transfer

Outline

As we reported in the last paragraph, *Sinorhizobium meliloti* is a model symbiont rhizobium able to establish symbiotic relationship with alfalfa. Its tripartite genomes comprise a chromosome and two megaplasmids, pSymA and pSymB, which carry genes related to the processes of symbiotic nitrogen fixation and growth in the rhizosphere, respectively. The horizontal gene transfer of symbiotic megaplasmid pSymA is a highly regulated by a system of 7 key genes named *rct* system. This system is repressed and activated following the perception of intrinsic and extrinsic signals that are still unknown and which lead to the repression of RctA. To date, to create new symbiont rhizobial strains able to infect new plant species is essential to discover which are the signals and the conditions which promote the mobilization of symbiosis plasmids. Systems biology approaches are crucial to face the challenge to synthetize in laboratory new PGPRs SynComs to be directly applied in field as biofertilizer in precision and sustainable agriculture.

To face the necessity to discover which are the signals that activate the plasmids mobilization *rct* system, we set out a new laboratory protocol for the in-vitro detection of the expression of the 7 key genes of the regulatory cascade. Briefly, each promoter of *rct* genes was cloned upstream to the green fluorescent protein (GFP) gene in the pOT2 plasmid. Recombinant pOT2 plasmids were individually transferred into *S. meliloti* and tested in different conditions. The number of cells expressing GFP was counted through spectrophotometer reading in preliminary tests and, consequently, through cytofluorimetric readings. These tests showed that the fraction of cells expressing GFP was around 10-4 over the total cells counted and one promoter was constitutively expressed. Further investigations through the recombinant panel of *S. meliloti* strains are necessary to discover and better understand the inducing conditions which allow the activation of the *rct* system and, consequently, the mobilization of symbiosis plasmids.

3.1 Horizontal gene transfer and conjugation process in rhizobia

From the results obtained in paper xiii and presented in the previous section, it emerged that the transfer of the symbiotic megaplasmid is not randomly occurring within the same clade. Moreover, previous work indicated that the phylogeny of the pSymA megaplamid in *S. meliloti* does not mirror chromosomal phylogeny[80,102]. It is then coming essential, even in the perspective of genome

manipulation of rhizobia[103] to understand when and how symbiotic megaplasmid move among strains.

Horizontal gene transfer (HGT), or lateral gene transfer, is clearly an essential process for the diversification and evolution of bacteria, which involves the movement of genetic information between organisms[104]. Referring to the evolutionary history of the rhizobium-legume symbiosis, the acquisition of symbiosis genes through HGT represents the key point for the origins of the major rhizobial clades and for the legume range expansion[105]. There are three main mechanisms to accomplish HGT between bacteria: transformation, transduction and conjugation. Bacterial conjugation is considered a highly efficient means of HGT[106]. The conjugation process involves transferring genetic material from one bacterium to another through direct cell-cell contact. During this process, one bacterium serves as the donor of the genetic material, and the other serves as the recipient. ssDNA molecules are transferred from donor to recipient bacterial cells. This HGT through conjugation depends on the presence of conjugative plasmids (CPs) or integrated conjugative elements (ICEs), which code for the essential genes for DNA transmission[107]. The rhizobial genera of *Sinorhizobium*, *Rhizobium*, *Cupriavidus* and *Paraburkholderia* carry the symbiotic genes on plasmids called pSyms. Other genera in the group, namely *Mesohizobium*, *Azorhizobium* and *Bradyrhizobium*, primarily have symbiotic genes on integrative and conjugative elements called ICEsyms.

While sym element mobility may seem beneficial, the process of HGT could lead negative effects for both bacterial donor and recipient. Compared to the majority of bacterial accessory traits, nodulation and nitrogen fixation processes involve the collaboration of a large suite of genes that modulate a complex series of events. Therefore, the mobility of symbiosis genes is, at first appearance, unexpected. First of all, conjugation requires the cell to invest in plasmid/ICE genome replication, conjugation pilus construction and the time required for transfer between host and recipient. During this time, the cell can become susceptible to phages which target the conjugative pilus. Secondly, for a successful transfer it is required that the recipient cell lacks a copy of the incoming element or an element sufficiently related to cause incompatibility, in the case of a plasmid. Moreover, some ICEs require integration sites that are unoccupied, and plasmids cannot coexist if their replication or partitioning systems are too closely related. Furthermore, symbiotic rhizobia strains must be able to do specific functions as to (i) respond to and communicate with their specific plant hosts, (ii) infect and form intracellular colonies within plant nodules (controlled by nod genes) and then (iii) undergo sophisticated cell differentiation in order to devote cellular metabolism to the highly energy-intensive process of nitrogen fixation. For these reasons, transferring the symbiosis cassette risks breaking up these collaborative genes and, in addition, to lose any previously beneficial adaptation acquired on the chromosome and other replicons[105]. On the other side, the transfer of symbiosis genes between bacteria seems to be an high expansive process in terms of time and energy to the bacterial donor[60] in itself and will also result in the creation of more new competitor symbionts for the donor bacteria for plant hosts in the microbial community. To date, these costs are often unpredictable[105].

The persistence of sym elements depends on reducing the costs imposed during transfer. However, the cost of plasmid acquisition is likely to be resolved through compensatory mutations. The success of transfer to a novel host will thus depend on the size of the initial cost, and the accessibility of compensatory mutations to relieve it.

Finally, the success of sym element transfer is also dependent on the integration and function of sym genes once acquired. Bacterial accessory genes prone to horizontal gene transfer are highly diverse, encoding functions, such as resistance traits, virulence factors or novel metabolic functions,

that are often made up of comparatively small operational units. Fortunately, the symbiosis cassette operates as a (relatively) self-contained operational unit. Genes that are heavily integrated into gene

networks are extremely costly to acquire as they are likely to lead to regulatory disruption. Consequently, accessory genes typically have a relatively low level of transcriptional connectivity[105].

3.2 Regulation mechanisms of the symbiotic megaplasmids conjugation process

The tight regulation of sym element transfer is one mechanism by which the costs of conjugation can be minimized. Furthermore, across replicons the capacity for mobilization is highly variable[105]. Many sym elements carry the genes required to initiate their own transfer via conjugation, while others depend on mobilization by other MGEs. To date, four major classes of conjugative machinery have been identified, classified and studied in the fast-growing rhizobia, i.e. Type I, Type II, Type III and Type IV. Type I and Type II conjugation systems were extensively studied [108,109], and Type III and Type IV conjugation systems were found recently in *Rhizobium leguminisarum* bv. *viciae* VF39SM[107,110].

— Type I: Quorum sensing (QS) mediated conjugation

Type I elements are generally (but not exclusively) regulated by Quorum sensing (QS). QS is a bacterial regulatory strategy dependent by population density and cell-cell signalling mechanism and allows bacteria to estimate their population size/density and to coordinate their behaviour and the expression of target genes using a small diffusible molecule. In rhizobia, the QS system usually depends on the molecule AHLs (N-acyl-homoserine lactones)[20,111,112]. Consequently, high population density induces conjugation, although the details of the regulatory networks that control these systems vary. Evidence suggests that QS via AHLs in more common among plant-associated bacteria belongs to genera such as *Agrobacterium*, *Rhizobium*, *Sinorhizobium*[105].

The best-studied QS mechanism is the one present in the Ti plasmid of Agrobacterium tumefaciens[113]. Conjugative transfer of the Ti plasmid involves two hierarchical signalling systems: the master system depends on the presence of opines. These molecules are secreted from the crown gall tumors induced on susceptible plants by A. tumefaciens and are able to bind to receptor proteins (such as AccR), and the complex allows the expression of TraR, which is the QS-dependant transcriptional regulator. The *traI*-encoded gene product is an acyl homoserine lactone (AHL) molecule involved in QS-regulation. When cellular density is low, the AHL molecules diffuse out of the cells and are rapidly diluted, while at high density population, the intracellular AHL molecules accumulate, and they can bind to a transcriptional regulator (TraR). This process activates the transcription of Dtr and Mpf genes involved in conjugal transfer, as well as positively autoregulate the *traI* and *traR* genes. Another protein, TraM, modulates the activity of TraR[20].

— Type II: RctA-mediated repression system

In type II elements, conjugation is under the control of RctA gene that constitutively represses transcription of the *virB* operon, involved for conjugation[109,114,115]. Very little is known about the environmental stimulus that alleviates RctA repression, suggesting that conjugation is limited to environments that are challenging to reproduce in the laboratory. The type II conjugation regulation has also been observed in the self-transmissible plasmid, pSymA, in *Sinorhizobium meliloti* strain 1021. The conjugation of pSymA in *S. meliloti* is repressed by the action of *rctA*, whose repression is antagonized by *rctB*. The environmental signal that triggers the expression of *rctB* to allow conjugal transfer has not been identified. Moreover, pSymA carries additional genes involved in the regulation of *rctB*, namely *rctR* and *rctC*. *rctR* encodes a protein of the GntR family of transcriptional regulators involved in the repression of *rctB*. *rctC* encodes a response regulator required for derepression of

conjugative functions through activating the *rctB* transcription. This type of conjugative plasmid and regulation system is now found relatively frequently in rhizobial genomes using bioinformatics searches[116]. However, the transfer process in these plasmids activates under optimal, as yet unknown environmental conditions involved in the inactivation of the *rctA* function[20,109,115].

— Type III elements are not able to self-mobilize as they lack genes required for mate pair formation. Although they are unable to initiate conjugation themselves, they can have retained the genes required for DNA transfer and replication[105].

— More recently, a fourth class of conjugative plasmids (type IV) has been identified, which uses a distinct repression pathway. This type of system is present on pSyms and non-sym plasmids, in a wide array of different rhizobial species[105].

Both the rhizobial symbionts that gain the functions and the plant hosts that depend on them can benefit greatly from this process, but control of conjugation rests predominantly with the MGEs that mediate *sym* genes transfer. Future work examining the evolutionary and ecological forces acting on these elements is therefore key to understanding the dynamics of this important symbiosis.

Decades of detailed work has revealed a complex and diverse ecosystem of MGEs within rhizobial genomes as well as a meticulous understanding of—at least some of—the diverse mechanisms that underlie this process. The recent discoveries of novel conjugation machineries among rhizobial plasmids demonstrates that this diversity is far from understood—but provides a firm grounding for future work applying ecological and evolutionary perspectives to this intracellular community[105].

3.2.1 RctA-mediated repression system in *Sinorhizobium meliloti*

Conjugation, or plasmid transfer, is an energy-consuming process, and therefore, these ssDNA transfer systems and regulatory networks maintain temporally and spatially controlled DNA transfer (*tra*) gene expression. By default, *tra* genes remain shut off in the cells which have conjugative elements and only become transfer competent when conditions are optimal. Even under optimal conditions, only a portion of the population undergoes this transition. It reduces the metabolic burden and the adverse effects of *tra* gene expression at the population level without affecting the potentiality of horizontal gene transfer. Because of this cooperative social behavior, conjugative elements can successfully thrive and evolve with bacteria to build a significant driving force of bacterial evolution. Indeed, conjugal transfer of symbiotic plasmids (pSyms) in rhizobia has played a major role in the evolution of these soil bacteria[117]. To date, the regulation of rhizobial plasmids conjugal transfer seems to involve at least two different transfer systems: quorum sensing (QS) regulated plasmids, where conjugal transfer responds positively to population density, and RctA-regulated plasmids which would only transfer under conditions that inactivate the transcription of the repressor RctA. To this class belongs for example the 1354 Kb pSymA of *Sinorhizobium meliloti* 1021[109]. *Sinorhizobium meliloti* establishes symbiotic relationship with alfalfa, *Medicago sativa*. As reported in the paragraph (2.4), its genome is composed by a chromosome and two megaplasmids which carry genes for diverse functions, including those related to symbiosis and nitrogen fixation. Indeed, on pSymA are localized most genes needed for nodulation and symbiotic nitrogen fixation, whereas pSymB harbours exopolysaccharide biosynthetic genes required for the establishment of symbiosis[115].

Thanks to recent studies conducted by Nogales et al., 2013, Blanca-Ordóñez et al., 2010 and Pérez-Mendoza et al., 2005, the RctA-mediated repression system has been clarified (Figure 11). They showed that pSymA carries an oriT and *traA1*, *traCDG*, and *virB1-virB11* operons encoding necessary

functions for self-transmissibility. On the other hand, for conjugal transfer, pSymB carries a *traA2* gene encoding a relaxase homologous to *traA1* from pSymA, and an oriT (origin of transfer) highly similar to the pSymA oriT sequence. The conjugative transfer of the *S. meliloti* pSymA (and of the mobilizable pSymB) is subject to rctA-dependent repression under laboratory conditions because *rctA* gene encodes a transcriptional repressor of *virB* and *tra* genes[109]. pSymB is a mobilizable plasmid and its conjugal transfer depends on certain pSymA functions like the *virB*-encoded type IV secretion system, and therefore is also subject to *rctA* repression. Thus, in *S. meliloti rctA* controls the conjugative transfer of more than 3 Mb of plasmid DNA[114]. RctA binds to specific DNA sequences of the *virB* operon promoter, suppressing the conjugal transfer. In addition, another gene, *rctB*, was identified dowstream of the *traA* gene. The overexpression of *rctB* from a heterologous promoter is able to antagonize the repressive effect of rctA on conjugal transfer[109], promoting transfer of symbiotic plasmid together with enhancing transcription of essential conjugal transfer genes. Furthermore, RctR inhibits pSymA conjugal transfer through repression of the *rctB* gene.

However, in *S. meliloti* the overexpression of *rctB* did not have an effect on *rctA* transcription, indicating that the antagonism of *rctB* on *rctA* gene could operate at the post-transcriptional level[109]. There are two additional regulatory elements in the *S. meliloti* pSymA that function upstream of *rctB* in the signalling cascade. On one hand is gene *rctC* (SMa0961), encoding a RR of the two components regulatory family. Our results indicate that expression of *rctC* is required for the transcriptional activation of *rctB* and therefore for derepression of the conjugative system.

This indicates that the key activation of *rctB* to antagonize repression by *rctA* may be achieved through different signalling cascades in different plasmids. Whether these differences determine that conjugation of different RctA type plasmids can proceed in response to different signals or environmental conditions is an issue that certainly deserves to be pursued.

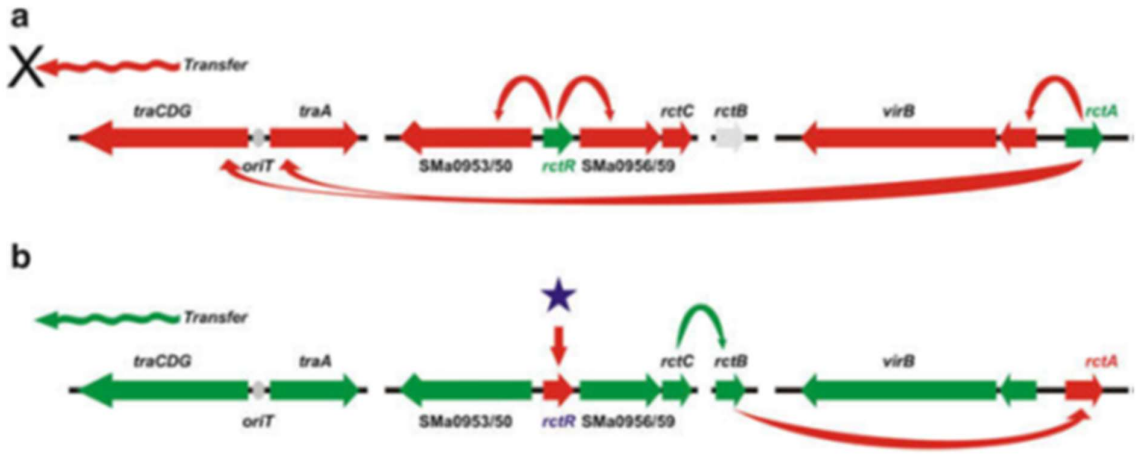


Figure 11. *Regulation of S. meliloti pSymA transfer through RctA repression. Genes whose transcription is repressed are in red, actively transcribed genes are in green. (a) Under standard conditions, constitutively expressed RctA represses transcription of the traA , traCDG , and virB operons, and RctR represses Sma 0953/50 and Sma 956/59/ rctC operons. (b) An unknown signal (blue star) causes repression of rctR , this allows expression of rctC , which induces rctB . RctB antagonizes RctA, allowing expression of Dtr and Mpf genes, leading to conjugative transfer (modifi ed from Nogales et al. 2013) required to activate rctB transcription and therefore plasmid conjugation. The rctR and rctC genes are absent in R. etli , indicating that although S. meliloti pSymA and R. etli pRet42a are both controlled by RctA and RctB regulators, different regulatory cascades modulate their expression[17].*

3.3 The question and the aims of our study

Recent studies demonstrated that the pSymA megaplasmid of *S. meliloti* is a self-conjugative plasmid but only in specific conditions which lead to the repression of RctA. Despite our knowledge about the *rct* mediated repressor system, it remains a future challenge the molecular mechanism that activate the *rct* regulatory cascade. The environmental conditions and the external and internal transduction signals involved in blocking RctA still remain unknown. To date, discovering which are the regulatory signals and the favourable conditions to mobilize the symbiosis plasmids is a prerogative to create new symbiont rhizobial strain able to infect new plant species. Through new sequencing technologies and new molecular approaches, it is indispensable to investigate on the regulatory cascade of megaplasmids. This challenge represents a powerful approach in systems biology to synthetize in laboratory PGPRs SynComs. This rhizobial communities could be applied, directly in field or indirectly with seeds coating, as biofertilizers in sustainable and precision agriculture to reduce the input of synthetic fertilizers.

Under these evidences and with the aim to better understand which are the signals that activate the plasmids mobilization *rct* system, we tested a new laboratory strategy to detect the expression of the seven genes of the *rct* regulatory cascade and we set out an original laboratory protocol to be applied for further investigations. In the following draft, we reported our preliminary tests and the protocol we used.

Section Work in progress:

II. Cangioli L., Fini M., Cambi G., Mengoni A., Fagorzi C. Study of the regulatory cascade involved in the megaplasmids conjugation in the model plant symbiotic bacterium Sinorhizobium meliloti by the creation of a GFP-based reporter system

Chapter 4

How epigenetics affects rhizobial strains differentiation and symbiosis?

Outline

Methylation of specific DNA sequences is involved in the modulation of genes expression in both prokaryotes and eukaryotes. In bacteria, it is well known that DNA methylation, in particular the presence of the methyl group in the N4 position of cytosine (N4-C) or the N6 position of adenine (N6-A), plays a role in different molecular mechanisms in immunity and regulation of cellular processes and these epigenetic traits are heritable. To date, we know that N6-methyladenine (6mA) is involved in chromosome replication, mismatch repair, conjugal transfer, antibiotic resistance and bacterial differentiation, whereas the N4-methyl-cytosine (4mC) is involved to pathogenesis. Hence, it seems to be clear that DNA methylation contributes to determination of relevant phenotypes, such as virulence and host colonization. Moreover, because of the different functions in which DNA methylation is involved, methylation patterns showed a great variability in many bacterial species.

In this chapter, we investigated on the different functions and variability of DNA methylation in bacteria. In particular, we wanted to verify whether DNA methylation system prevents the horizontal gene transfer of mobile genetic elements despite it has been observed that sometimes the acquisition of new plasmids occurs, and whether the accumulation of DNA methylation in *S. meliloti* specific sites (GANTC) is involved in cell cycle regulation. Finally, we developed a computational tool, named MeStudio, that maps methylated DNA motifs with respect to coding and noncoding sequences in the genome and allowed us to do a population epigenomic analysis of *S. meliloti* strains, analysing the strain-by-strain and intragenomic variation related to multireplicon genome structure of *S. meliloti* and presence and type of coding sequences.

From these studies, we observed that DNA methylation has different occurrences among strains, between pSymA and the rest of the genome, and, finally, between intergenic and coding sequences. Methylation occurrences differ also during bacteroid differentiation, in particular the GANTC motif that showed differential proportion of methylated sites along pSymA in bacteroid compared to free living cells.

4.1 DNA methylation in bacteria

Epigenetic information is essential to modulate the expression levels of important genes in both prokaryotes and eukaryotes. These epigenetic traits are heritable changes in gene expression without modifying the DNA sequences itself[119]. The universal mechanism of epigenetic regulation is the DNA methylation, which is the process that adds methyl groups to DNA nucleotides by the enzymes DNA methyltransferase[120].

The improved genome-scale mapping of methylation allowed as to discover that the function of DNA methylation seems to vary with context. In eukaryotic organisms epigenetic is most extensively studied and DNA methylation seems to be involved in the control of some pivotal biological processes, such as cellular differentiation, development and pathogenesis[119]. On the other side, in prokaryotic microorganisms, epigenetics is known to be involved in different molecular mechanisms such as DNA repair, DNA-protein interactions, gene expression, cell cycle progression and self-DNA recognition (the Restriction-Modification systems)[121,122].
In bacteria, the predominant DNA methyltransferases are the exocyclic amino methyltransferases that transfer a methyl group to the N4 position of cytosine (N4-C) or the N6 position of adenine (N6-A), e.g., Dam and CcrM[120]. Recent studies demonstrated that N6-methyladenine (6mA) is involved in many events, such as chromosome replication, mismatch repair, conjugal transfer, antibiotic resistance[121] and bacterial differentiation[123], whereas the N4-methyl- cytosine (4mC) is involved to pathogenesis[118]. It is now evident that DNA methylation in bacteria plays important roles in affecting relevant phenotypes such as virulence and host colonization[120].

Principally, there are two main categories of DNA methyltransferases: methyltransferases in Restriction-Modification systems ("R-M systems") and "solitary" or "orphan" methyltransferases. In R-M systems, DNA methyltransferases (MTases) are associated with restriction enzymes (REases). In most bacteria, the R-M systems act as immune response, protecting their own DNA while degrading foreign DNA. In addition, MTases from R-M systems have been found to be involved in epigenetic regulation in some bacteria. Consequently, R-M systems are thought to be essential for host colonization by pathogenic bacteria[120].
The second category of methyltransferases, the "solitary" or "orphan" methylases, exists with no association with any restriction enzymes. MTases from R-M systems cannot cleave unmethylated foreign DNA. The origin of orphan methyltransferases have been attributed to the horizontal gene transfer, with only the MTase part of an R-M system being transferred[120].

4.2 Unravelling DNA methylation role in genomes

Concerning the methylation patterns, it has been observed a great variability in many bacterial species. This variability is related to the different functions that the methylation system carries out. Among all these functions, the most studied in prokaryotes are those related to the R-M system, the adaptation to diverse environmental conditions, and to virulence and pathogenicity [118,120].

Since R-M systems are able to distinguish self-DNA and foreign DNA, it is evident that DNA methyltransferases play a conservative function of the bacterial genome during evolution, preventing the horizontal gene transfer (HGT). Despite this strategy to prevent attacks from external DNA, it has been observed that the acquisition of new mobile genetic elements (MGEs) can occur. Additionally, since evidences for a role of DNA methylation in regulating bacterial gene expression are accumulating and in *S. meliloti* methylation of DNA at specific sites (GANTC) is pivotal for cell cycle regulation[124,125], we wondered if DNA methylation may be related to bacteroid differentiation during symbiotic interaction (xi) and we developed a computational tool (xii) to map methylated DNA motifs with respect to coding and noncoding sequences in the genome. Finally, this tool, named MeStudio, was used in a population epigenomic analysis of *S. meliloti* strains, investigating strain-by-strain variation and intragenomic variation related to multireplicon genome structure of *S. meliloti* and presence and type of coding sequences (xiii).

xi) diCenzo, G. C., Cangioli, L., Nicoud, Q., Cheng, J. H., Blow, M. J., Shapiro, N., ... & Mergaert, P. (2022). DNA methylation in Ensifer species during free-living growth and during nitrogen-fixing symbiosis with Medicago spp. *Msystems*, *7*(1), e01092-21. doi: https://doi.org/10.1128/mSystems.01092-21

xii) Riccardi, C., Passeri, I., Cangioli, L., Fagorzi, C., Fondi, M., Mengoni, A. (2023). Crossing Bacterial Genomic Features and Methylation Patterns with MeStudio: An Epigenomic Analysis Tool. *International Journal of Molecular Sciences.* doi: https://doi.org/10.3390/ijms24010159

xiii) Passeri, I., Cangioli, L., Fondi, M., Mengoni, A., Fagorzi, C. (2023). The pan-epigenome of the symbiotic nitrogen fixing bacterium Sinorhizobium meliloti unravels unexpected variability of DNA-methylation profiles in closely related strains. bioRxiv 2023.05.31.542881; doi: https://doi.org/10.1101/2023.05.31.542881

Results from these studies showed the presence of a large number of methylated motifs which differ in presence and fraction of methylated sites among strains and between pSymA and the rest of the genomes, again emphasizing the nature of pSymA as alien element in the *S. meliloti* genome. Moreover, we observed that methylation occur differentially between part of intergenic sequences and coding sequences, strongly suggesting a role of several methylated DNA motifs in regulation of gene expression. This modulation of methylated motifs was also observed during bacteroid differentiation, where the GANTC motif, related to CcrM methylation (the DNA MTase controlling *S. meliloti* cell cycle) shows a differential proportion of methylated sites along pSymA in bacteroid compared to free living cells, suggesting a dysregulation of cell cycle during symbiosis in relation to DNA methylation.

Perspectives

In this three-years journey, the word around which my thesis was developed is *sustainability*. The Technical Advisory Committee of Consultative Group on International Agricultural Research defined sustainability as "the successful management of resources for agriculture to satisfy changing human needs while maintaining or enhancing the quality of the environment and conserving resources". Indeed, sustainable agriculture aims to improve the resource efficiency and conservation while maintaining an environment favourable for evolution of all species. What is important to keep in mind is the attention that the sustainability theories reserve to future generations. This means that sustainability is not only a matter of environment but also wellness and economy.

Nowadays, the intensive use of chemical fertilizers and pesticides has led to the deterioration of the quality of both the soil as well as the environment[126]. On the other hand, these agrochemical products have been (and they are still) important to satisfy the increasing demand for food, feed, and bioenergy. Recent studies estimate that agricultural production must increase by about 70% from its current level to meet the demand for a growing population that will reach 9.8 billion people by 2050.
A very innovative and ecofriendly strategy to face several agricultural challenges due to climate change, limited resources, and land degradation is represented by the naturally occurring microbial systems[31].

The development of sustainable agriculture practices that harness crop-associated microbiomes represents a promising approach to either increase or sustain higher yields while maintaining overall soil health and fertility[126]. Recent advances in 'omics technologies and the innovative computational tools allowed us to systematically explore and discover novel gene functions in microbiomes. Biocontrol, biofertilization, and biostimulation can now be engineered[31].

In many cases, microbial inoculant products that contain single species either fail to colonise plant in field conditions or do not provide expected benefits. Conversely, it is increasingly recognized that the use of synthetic microbial communities (SynComs) can overcome some of these challenges[28]. SynComs include multiple microbial species which can grow together and provide collective benefits to the plant host. Synergistic interactions between compatible but functionally diverse microbial strains can lead to better adaptation to new environments and simultaneously offer a broader range of plant growth-promoting functional traits to improve plant phenotypes and ultimately enhance crop productivity.

However, our understanding of biotic mechanisms that underpin positive impacts on plant fitness such as improved plant performance and the effects of delivery mode of synthetic community are poorly understood. Timing and mode of application can also have impacts on microbial inoculant efficacy. Despite potential benefits, these are critical knowledge gaps that constrain field efficacy of SynCom and hence largescale adoption by the farming community.

From our point of view, the idea of exploiting plant microbiome engineering to improve the sustainability of agriculture represents a realistic starting point to fulfil the needs of the present without compromising the needs of the future. A more holistic understanding is still needed to better understand the intermicrobial interactions within the microbiota of plants and to better define the functional relevance of the microbial networks for holobiont fitness[29,127]. Prokaryotic and eukaryotic microbes have evolved a myriad of cooperative and competitive interaction mechanisms that shape

and likely stabilize microbial assemblages on plant tissues. However, most of the data are derived from one-to-one interaction studies, and only few incorporate complex microbial communities in controlled laboratory conditions to reconstitute the plant microbiota and to understand the role of intermicrobial interactions. Such experiments will shed new light on the fundamental principles that govern the assembly of complex microbial communities and the maintenance of host-microbial homeostasis. It is important to consider microbe-microbe interactions to accept or reject the hologenome theory, which postulates that selection can operate on horizontally acquired plant microbiota members. According to this concept, it is likely that microbes that tightly associate with plants also evolve community level microbe-microbe interaction strategies that allow them to persist within the plant holobiont[128,129].

The complexity of small things is sometimes surprising, especially when looking at the positive effects a bacterium can bring to the growth of host (i.e. a plant). The idea that the presence in the soil of bacterial strains able to ameliorate the production of a crop is intriguing; the possibility to use specific bacterial strains as if they are “precision energy supplements” for specific plant species gives the idea of how powerful and rich of novelty research in the field of sustainable agriculture is becoming.

A challenge is to make people aware that the use of bacterial inoculants in agriculture is useful for limiting the use of synthetic fertilizers. The challenge in the challenge is to understand the rules (mechanisms) of the plant-association game and use our biotechnological skills to safely combine multiple traits of different bacterial strains to obtain more efficient inoculants. Omics technologies are providing huge amounts of data that help us to deeply understand the finest mechanisms of bacterial metabolism and the precise composition of microbial communities present in the most disparate environments. A lot of work has still to be carried out to answer the many questions on our small, simple but indeed complex (micro)organisms. We have a lot of inventiveness and many resources to be brought into play.

Bibliography

1. Berg, G., Grube, M., Schloter, M. & Smalla, K. Unraveling the plant microbiome: looking back and future perspectives. *Front. Microbiol.* **5**, (2014).
2. Afridi, M. S. *et al.* Plant Microbiome Engineering: Hopes or Hypes. *Biology* **11**, 1782 (2022).
3. Santos, L. F. & Olivares, F. L. Plant microbiome structure and benefits for sustainable agriculture. *Curr. Plant Biol.* **26**, 100198 (2021).
4. Bano, S., Wu, X. & Zhang, X. Towards sustainable agriculture: rhizosphere microbiome engineering. *Appl. Microbiol. Biotechnol.* **105**, 7141–7160 (2021).
5. Orozco-Mosqueda, M. del C., Rocha-Granados, M. del C., Glick, B. R. & Santoyo, G. Microbiome engineering to improve biocontrol and plant growth-promoting mechanisms. *Microbiol. Res.* **208**, 25–31 (2018).
6. Sessitsch, A., Pfaffenbichler, N. & Mitter, B. Microbiome Applications from Lab to Field: Facing Complexity. *Trends in Plant Science* Preprint at https://doi.org/10.1016/j.tplants.2018.12.004 (2019).
7. Choi, K., Khan, R. & Lee, S.-W. Dissection of plant microbiota and plant-microbiome interactions. *J. Microbiol.* **59**, 281–291 (2021).
8. Bulgarelli, D. *et al.* Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* **488**, 91–95 (2012).
9. Oluwafunto Ajayi, O. & Dianda, M. Plants and Their Microbes. in *Symbiosis in Nature* (ed. Cid Rigobelo, E.) (IntechOpen, 2023). doi:10.5772/intechopen.109917.
10. Xu, L. *et al.* Holo-omics for deciphering plant-microbiome interactions. *Microbiome* **9**, 69 (2021).
11. Trivedi, P., Leach, J. E., Tringe, S. G., Sa, T. & Singh, B. K. Plant–microbiome interactions: from community assembly to plant health. *Nat. Rev. Microbiol.* **18**, 607–621 (2020).
12. Zhang, J. *et al.* Harnessing the plant microbiome to promote the growth of agricultural crops. *Microbiol. Res.* **245**, 126690 (2021).
13. Lyu, D. *et al.* Plant Holobiont Theory: The Phytomicrobiome Plays a Central Role in Evolution and Success. *Microorganisms* **9**, 675 (2021).
14. Afridi, M. S. *et al.* Harnessing microbial multitrophic interactions for rhizosphere microbiome engineering. *Microbiol. Res.* **265**, 127199 (2022).
15. He, M., He, C.-Q. & Ding, N.-Z. Abiotic Stresses: General Defenses of Land Plants and Chances for Engineering Multistress Tolerance. *Front. Plant Sci.* **9**, 1771 (2018).
16. Hiltner, L. Über neuere Erfahrungen und Probleme auf dem Gebiet der Bodenbakteriologie und unter besonderer Berücksichtigung der Gründüngung und Brache. *Arb. Dtsch. Landwirtsch. Ges.* **98**, 59 (1904).
17. Bakker, M. G., Schlatter, D. C., Otto-Hanson, L. & Kinkel, L. L. Diffuse symbioses: roles of plant–plant, plant–microbe and microbe–microbe interactions in structuring the soil microbiome. *Mol. Ecol.* **23**, 1571–1583 (2014).
18. Munir, N. *et al.* Mechanisms and Strategies of Plant Microbiome Interactions to Mitigate Abiotic Stresses. *Agronomy* **12**, 2069 (2022).
19. Abd-Alla, M. H., Al-Amri, S. M. & El-Enany, A.-W. E. Enhancing Rhizobium–Legume Symbiosis and Reducing Nitrogen Fertilizer Use Are Potential Options for Mitigating Climate Change. *Agriculture* **13**, 2092 (2023).

20. Brom, S., Pistorio, M., Romero, D. & Torres-Tejerizo, G. Boundaries for Conjugative Transfer of Rhizobial Plasmids: Restraining and Releasing Factors. in *Plasticity in Plant-Growth-Promoting and Phytopathogenic Bacteria* (ed. Katsy, E. I.) 43–54 (Springer New York, New York, NY, 2014). doi:10.1007/978-1-4614-9203-0_3.
21. Martins, S. J. *et al.* The Use of Synthetic Microbial Communities to Improve Plant Health. *Phytopathology®* **113**, 1369–1379 (2023).
22. Jia, H.-M., Li, B., Wu, Y.-R., Ma, Y.-T. & Yan, Z.-Y. The construction of synthetic communities improved the yield and quality of Salvia miltiorrhiza Bge. *J. Appl. Res. Med. Aromat. Plants* **34**, 100462 (2023).
23. Wang, C. *et al.* Functional assembly of root-associated microbial consortia improves nutrient efficiency and yield in soybean. *J. Integr. Plant Biol.* **63**, 1021–1035 (2021).
24. De la Vega-Camarillo, E. *et al.* Promotion of the growth and yield of Zea mays by synthetic microbial communities from Jala maize. *Front. Microbiol.* **14**, 1167839 (2023).
25. Kaur, S. *et al.* Synthetic community improves crop performance and alters rhizosphere microbial communities. *J. Sustain. Agric. Environ.* **1**, 118–131 (2022).
26. Izawa, T. Reloading DNA History in Rice Domestication. *Plant Cell Physiol.* **63**, 1529–1539 (2022).
27. Nerva, L. *et al.* Breeding toward improved ecological plant–microbiome interactions. *Trends Plant Sci.* **27**, 1134–1143 (2022).
28. Trivedi, P., Batista, B. D., Bazany, K. E. & Singh, B. K. Plant–microbiome interactions under a changing world: responses, consequences and perspectives. *New Phytol.* **234**, 1951–1959 (2022).
29. Hacquard, S., Wang, E., Slater, H. & Martin, F. Impact of global change on the plant microbiome. *New Phytol.* **234**, 1907–1909 (2022).
30. Vimal, S. R., Singh, J. S., Kumar, A. & Prasad, S. M. Plant genotype-microbiome engineering as nature-based solution (NbS) for regeneration of stressed agriculture: A review. *Sci. Hortic.* **321**, 112258 (2023).
31. Vaccaro, F., Cangioli, L., Mengoni, A. & Fagorzi, C. Synthetic plant microbiota challenges in nonmodel species. *Trends Microbiol.* **30**, 922–924 (2022).
32. Lau, S.-E., Teo, W. F. A., Teoh, E. Y. & Tan, B. C. Microbiome engineering and plant biostimulants for sustainable crop improvement and mitigation of biotic and abiotic stresses. *Discov. Food* **2**, 9 (2022).
33. Afridi, M. S. *et al.* New opportunities in plant microbiome engineering for increasing agricultural sustainability under stressful conditions. *Front. Plant Sci.* **13**, 899464 (2022).
34. Jansson, J. K., McClure, R. & Egbert, R. G. Soil microbiome engineering for sustainability in a changing environment. *Nat. Biotechnol.* (2023) doi:10.1038/s41587-023-01932-3.
35. Dubey, S. & Sharma, S. Rhizospheric Engineering by Plant-Mediated Indirect Selection of Microbiome for Agricultural Sustainability. *Crit. Rev. Plant Sci.* **40**, 379–397 (2021).
36. Phour, M. & Sindhu, S. S. Mitigating abiotic stress: microbiome engineering for improving agricultural production and environmental sustainability. *Planta* **256**, 85 (2022).
37. Busby, P. E. *et al.* Research priorities for harnessing plant microbiomes in sustainable agriculture. *PLOS Biol.* **15**, e2001793 (2017).
38. Brahim, A. H. *et al. Seed-Biopriming of Durum Wheat with Diazotrophic Plant Growth Promoting Bacteria (PGPB) Enhanced Tolerance to Fusarium Head Blight (FHB) and Salinity Stress*. https://www.researchsquare.com/article/rs-7274/v2 (2019) doi:10.21203/rs.2.16636/v2.
39. Ngalimat, M. S. *et al.* Plant Growth-Promoting Bacteria as an Emerging Tool to Manage Bacterial Rice Pathogens. *Microorganisms* **9**, 682 (2021).

40. Yang, S. *et al.* Emerging Pathways for Engineering the Rhizosphere Microbiome for Optimal Plant Health. *J. Agric. Food Chem.* **71**, 4441–4449 (2023).
41. Kalam, S., Basu, A. & Podile, A. R. Difficult-to-culture bacteria in the rhizosphere: The underexplored signature microbial groups. *Pedosphere* **32**, 75–89 (2022).
42. Rodriguez, P. A. *et al.* Systems Biology of Plant-Microbiome Interactions. *Mol. Plant* **12**, 804–821 (2019).
43. Dubey, S. & Sharma, S. Rhizospheric Microbiome Engineering as a Sustainable Tool in Agriculture: Approaches and Challenges. in *Microbial Diversity in Ecosystem Sustainability and Biotechnological Applications: Volume 2. Soil & Agroecosystems* (eds. Satyanarayana, T., Das, S. K. & Johri, B. N.) 257–272 (Springer Singapore, Singapore, 2019). doi:10.1007/978-981-13-8487-5_11.
44. Tilman, D. Global environmental impacts of agricultural expansion: The need for sustainable and efficient practices. *Proc. Natl. Acad. Sci.* **96**, 5995–6000 (1999).
45. Smith, B. E. Nitrogenase reveals its inner secrets. *Science* Preprint at https://doi.org/10.1126/science.1076659 (2002).
46. Erisman, J. W., Galloway, J., Seitzinger, S., Bleeker, A. & Butterbach-Bahl, K. Reactive nitrogen in the environment and its effect on climate change. *Curr. Opin. Environ. Sustain.* **3**, 281–290 (2011).
47. Duinen, G. V. How a century of ammonia synthesis changed the world. *Nat. Geosci.* **1**, (2008).
48. Vance, C. P. Enhanced Agricultural Sustainability Through Biological Nitrogen Fixation. in *Biological Fixation of Nitrogen for Ecology and Sustainable Agriculture* (eds. Legocki, A., Bothe, H. & Pühler, A.) 179–186 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1997).
49. Saha, B. *et al.* Biological Nitrogen Fixation for Sustainable Agriculture. in *Agriculturally Important Microbes for Sustainable Agriculture: Volume 2: Applications in Crop Production and Protection* (eds. Meena, V. S., Mishra, P. K., Bisht, J. K. & Pattanayak, A.) 81–128 (Springer Singapore, Singapore, 2017). doi:10.1007/978-981-10-5343-6_4.
50. Postgate, J. R. The fundamentals of nitrogen fixation. in (1978).
51. Shridhar, B. S. Nitrogen Fixing Microorganisms. (2012).
52. Vadakattu, G. & Paterson, J. Free-living bacteria lift soil nitrogen supply. *Farming Ahead* **169**, (2006).
53. Herridge, D. F., Peoples, M. B. & Boddey, R. M. Global inputs of biological nitrogen fixation in agricultural systems. *Plant and Soil* Preprint at https://doi.org/10.1007/s11104-008-9668-3 (2008).
54. Paffetti, D. *et al.* Influence of plant genotype on the selection of nodulating Sinorhizobium meliloti strains by Medicago sativa.
55. Finzi, A. C. *et al.* Rhizosphere processes are quantitatively important components of terrestrial carbon and nutrient cycles. *Glob. Change Biol.* **21**, (2015).
56. Vitousek, P. M., Menge, D. N. L., Reed, S. C. & Cleveland, C. C. Biological nitrogen fixation: rates, patterns and ecological controls in terrestrial ecosystems. *Philos. Trans. R. Soc. B Biol. Sci.* **368**, 20130119 (2013).
57. PEOPLES, M. B. & CRASWELL, E. T. Biological nitrogen fixation: Investments, expectations and actual contributions to agriculture. *Plant Soil* **141**, 13–39 (1992).
58. Zahran, H. Rhizobia from wild legumes: diversity, taxonomy, ecology, nitrogen fixation and biotechnology. *J. Biotechnol.* **91**, 143–153 (2001).
59. Vance, C. P. Symbiotic Nitrogen Fixation and Phosphorus Acquisition. Plant Nutrition in a World of Declining Renewable Resources. *Plant Physiol.* **127**, 390–397 (2001).

60. Herridge, D. F., Peoples, M. B. & Boddey, R. M. Global inputs of biological nitrogen fixation in agricultural systems. *Plant Soil* **311**, 1–18 (2008).
61. Smil, V. *Enriching the Earth: Fritz Haber, Carl Bosch, and the Transformation of World Food Production*. (The MIT Press, 2000). doi:10.7551/mitpress/2767.001.0001.
62. Checcucci, A., DiCenzo, G. C., Bazzicalupo, M. & Mengoni, A. Trade, Diplomacy, and Warfare: The Quest for Elite Rhizobia Inoculant Strains. *Front. Microbiol.* **8**, 2207 (2017).
63. West, S. A., Kiers, E. T., Simms, E. L. & Denison, R. F. Sanctions and mutualism stability: why do rhizobia fix nitrogen? *Proc. R. Soc. Lond. B Biol. Sci.* **269**, 685–694 (2002).
64. Islam, S., A. Aziz, M., M. Dar, Z. & Masood, A. Plant Growth Promoting Rhizobacterial Consortium: A Sustainable Crop Production Strategy. in *Symbiosis in Nature* (ed. Cid Rigobelo, E.) (IntechOpen, 2023). doi:10.5772/intechopen.109913.
65. Checcucci, A. *et al.* Role and Regulation of ACC Deaminase Gene in Sinorhizobium meliloti: Is It a Symbiotic, Rhizospheric or Endophytic Gene? *Front. Genet.* **8**, (2017).
66. Li, Z., Ma, Z., Hao, X., Rensing, C. & Wei, G. Genes Conferring Copper Resistance in *Sinorhizobium meliloti* CCNWSX0020 Also Promote the Growth of *Medicago lupulina* in Copper-Contaminated Soil. *Appl. Environ. Microbiol.* **80**, 1961–1971 (2014).
67. Lu, M. *et al.* Transcriptome Response to Heavy Metals in Sinorhizobium meliloti CCNWSX0020 Reveals New Metal Resistance Determinants That Also Promote Bioremediation by Medicago lupulina in Metal-Contaminated Soil. *Appl. Environ. Microbiol.* **83**, e01244-17 (2017).
68. Wani, P. A., Khan, M. S. & Zaidi, A. Effect of metal tolerant plant growth promoting Bradyrhizobium sp. (vigna) on growth, symbiosis, seed yield and metal uptake by greengram plants. *Chemosphere* **70**, 36–45 (2007).
69. Bulgarelli, D., Schlaeppi, K., Spaepen, S., Van Themaat, E. V. L. & Schulze-Lefert, P. Structure and Functions of the Bacterial Microbiota of Plants. *Annu. Rev. Plant Biol.* **64**, 807–838 (2013).
70. Sprent, J. I., Ardley, J. & James, E. K. Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytologist* Preprint at https://doi.org/10.1111/nph.14474 (2017).
71. Checcucci, A., diCenzo, G. C., Perrin, E., Bazzicalupo, M. & Mengoni, A. Chapter 3 - Genomic Diversity and Evolution of Rhizobia. in *Microbial Diversity in the Genomic Era* (eds. Das, S. & Dash, H. R.) 37–46 (Academic Press, 2019). doi:10.1016/B978-0-12-814849-5.00003-4.
72. Estrada-de los Santos, P. *et al.* Trinickia dabaoshanensis sp. nov., a new name for a lost species. *Arch. Microbiol.* (2019) doi:10.1007/s00203-019-01703-2.
73. Martens, M. *et al.* Multilocus sequence analysis of Ensifer and related taxa. *Int. J. Syst. Evol. Microbiol.* **57**, 489–503 (2007).
74. Denison, R. F. & Kiers, E. T. Lifestyle alternatives for rhizobia: mutualism, parasitism, and forgoing symbiosis. *FEMS Microbiol. Lett.* **237**, 187–193 (2004).
75. Gage, D. J. Infection and Invasion of Roots by Symbiotic, Nitrogen-Fixing Rhizobia during Nodulation of Temperate Legumes. *Microbiol. Mol. Biol. Rev.* **68**, 280–300 (2004).
76. Capela, D. *et al.* Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. *Proc. Natl. Acad. Sci.* **98**, 9877–9882 (2001).
77. Checcucci, A. Plant colonization: exploring the genetic basis of rhizobial symbiotic performance and bacterial invasion.
78. Honeycutt, R. J., McCLELLAND, M. & Sobral, B. W. Physical map of the genome of Rhizobium meliloti 1021. *J. Bacteriol.* **175**, 6945–6952 (1993).
79. Galibert, F. *et al.* The Composite Genome of the Legume Symbiont *Sinorhizobium meliloti*. *Science* **293**, 668–672 (2001).

80. Galardini, M., Pini, F., Bazzicalupo, M., Biondi, E. G. & Mengoni, A. Replicon-dependent bacterial genome evolution: the case of Sinorhizobium meliloti. *Genome Biol. Evol.* **5**, 542–558 (2013).
81. Laranjo, M., Alexandre, A. & Oliveira, S. Legume growth-promoting rhizobia: An overview on the Mesorhizobium genus. *Spec. Issue Plant Growth Promot.* **169**, 2–17 (2014).
82. Oke, V. & Long, S. R. Bacterial genes induced within the nodule during the *Rhizobium* –legume symbiosis. *Mol. Microbiol.* **32**, 837–849 (1999).
83. Wang, D., Yang, S., Tang, F. & Zhu, H. Symbiosis specificity in the legume - rhizobial mutualism: Host specificity in root nodule symbiosis. *Cell. Microbiol.* **14**, 334–342 (2012).
84. Brewin, N. J. Plant Cell Wall Remodelling in the Rhizobium–Legume Symbiosis. *Crit. Rev. Plant Sci.* **23**, 293–316 (2004).
85. Turgeon, B. G. & Bauer, W. D. Ultrastructure of infection-thread development during the infection of soybean by Rhizobium japonicum. *Planta* **163**, 328–349 (1985).
86. Deakin, W. J. & Broughton, W. J. Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nat. Rev. Microbiol.* **7**, 312–320 (2009).
87. Smit, G., Swart, S., Lugtenberg, B. J. J. & Kijne, J. W. Molecular mechanisms of attachment of Rhizobium bacteria to plant roots. *Mol. Microbiol.* **6**, 2897–2903 (1992).
88. Kondorosi, E., Banfalvi, Z. & Kondorosi, A. Physical and genetic analysis of a symbiotic region of Rhizobium meliloti: Identification of nodulation genes. *Mol. Gen. Genet. MGG* **193**, 445–452 (1984).
89. Schultze, M. & Kondorosi, A. REGULATION OF SYMBIOTIC ROOT NODULE DEVELOPMENT. *Annu. Rev. Genet.* **32**, 33–57 (1998).
90. David, M. *et al.* Cascade regulation of nif gene expression in Rhizobium meliloti. *Cell* **54**, 671–683 (1988).
91. Carelli, M. *et al.* Genetic Diversity and Dynamics of *Sinorhizobium meliloti* Populations Nodulating Different Alfalfa Cultivars in Italian Soils. *Appl. Environ. Microbiol.* **66**, 4785–4789 (2000).
92. Heath, K. D. & Tiffin, P. STABILIZING MECHANISMS IN A LEGUME-RHIZOBIUM MUTUALISM. *Evolution* **63**, 652–662 (2009).
93. Heath, K. D., Stock, A. J. & Stinchcombe, J. R. Mutualism variation in the nodulation response to nitrate. *J. Evol. Biol.* **23**, 2494–2500 (2010).
94. diCenzo, G. C. *et al.* Multidisciplinary approaches for studying rhizobium–legume symbioses. *Can. J. Microbiol.* **65**, 1–33 (2019).
95. Moura, G. G. D. de, Remigi, P., Masson-Boivin, C. & Capela, D. Experimental Evolution of Legume Symbionts: What Have We Learnt? *Genes* **11**, (2020).
96. Marchetti, M. *et al.* Shaping Bacterial Symbiosis With Legumes by Experimental Evolution. *Mol. Plant-Microbe Interactions®* **27**, 956–964 (2014).
97. Pini, F., Galardini, M., Bazzicalupo, M. & Mengoni, A. Plant-Bacteria Association and Symbiosis: Are There Common Genomic Traits in Alphaproteobacteria? *Genes* **2**, 1017–1032 (2011).
98. Geddes, B. A., Kearsley, J. V. S., Morton, R. A., diCenzo, G. C. & Finan, T. M. The genomes of rhizobia. *Adv. Bot. Res.* (2020).
99. Pérez Carrascal, O. M. *et al.* Population genomics of the symbiotic plasmids of sympatric nitrogen-fixing Rhizobium species associated with Phaseolus vulgaris. *Environ. Microbiol.* **18**, 2660–2676 (2016).
100. diCenzo, G. C. & Finan, T. M. The Divided Bacterial Genome: Structure, Function, and Evolution. *Microbiol. Mol. Biol. Rev.* **81**, e00019-17 (2017).

101. Kuzmanović, N., Fagorzi, C., Mengoni, A., Lassalle, F. & diCenzo, G. C. Taxonomy of Rhizobiaceae revisited: proposal of a new framework for genus delimitation. *Int. J. Syst. Evol. Microbiol.* **72**, (2022).
102. Riley, A. B., Grillo, M. A., Epstein, B., Tiffin, P. & Heath, K. D. Discordant population structure among rhizobium divided genomes and their legume hosts. *Mol. Ecol.* **32**, 2646–2659 (2023).
103. Checcucci, A. *et al.* Creation and Characterization of a Genomically Hybrid Strain in the Nitrogen-Fixing Symbiotic Bacterium *Sinorhizobium meliloti*. *ACS Synth. Biol.* **7**, 2365–2378 (2018).
104. Burmeister, A. R. Horizontal gene transfer. *Evol. Med. Public Health* **2015**, 193–194 (2015).
105. Wardell, G. E., Hynes, M. F., Young, P. J. & Harrison, E. Why are rhizobial symbiosis genes mobile? *Philos. Trans. R. Soc. B Biol. Sci.* **377**, 20200471 (2022).
106. Bañuelos-Vazquez, L. A., Torres Tejerizo, G. & Brom, S. Regulation of conjugative transfer of plasmids and integrative conjugative elements. *SI ISPB Plasmid 2016* **91**, 82–89 (2017).
107. Koraimann, G. & Wagner, M. A. Social behavior and decision making in bacterial conjugation. *Front. Cell. Infect. Microbiol.* **4**, (2014).
108. Danino, V. E., Wilkinson, A., Edwards, A. & Downie, J. A. Recipient-induced transfer of the symbiotic plasmid pRL1JI in Rhizobium leguminosarum bv. viciae is regulated by a quorum-sensing relay. *Mol. Microbiol.* **50**, 511–525 (2003).
109. Pérez-Mendoza, D. *et al.* Identification of the *rctA* Gene, Which Is Required for Repression of Conjugative Transfer of Rhizobial Symbiotic Megaplasmids. *J. Bacteriol.* **187**, 7341–7350 (2005).
110. Ding, H., Yip, C. B. & Hynes, M. F. Genetic Characterization of a Novel Rhizobial Plasmid Conjugation System in Rhizobium leguminosarum bv. viciae Strain VF39SM. *J. Bacteriol.* **195**, 328–339 (2013).
111. Acosta-Jurado, S., Fuentes-Romero, F., Ruiz-Sainz, J.-E., Janczarek, M. & Vinardell, J.-M. Rhizobial exopolysaccharides: Genetic regulation of their synthesis and relevance in symbiosis with legumes. *Int. J. Mol. Sci.* **22**, 6233 (2021).
112. Achchi, H. G. & Hemananda, K. M. Genomic and Proteomic Approaches to Understanding Conjugation in Rhizobium leguminosarum VF39. (2023).
113. Cha, C., Gao, P., Chen, Y.-C., Shaw, P. D. & Farrand, S. K. Production of Acyl-Homoserine Lactone Quorum-Sensing Signals by Gram-Negative Plant-Associated Bacteria. *Mol. Plant-Microbe Interactions®* **11**, 1119–1129 (1998).
114. Blanca-Ordóñez, H. *et al.* pSymA-Dependent Mobilization of the *Sinorhizobium meliloti* pSymB Megaplasmid. *J. Bacteriol.* **192**, 6309–6312 (2010).
115. Nogales, J., Blanca-Ordóñez, H., Olivares, J. & Sanjuán, J. Conjugal transfer of the *S inorhizobium meliloti* 1021 symbiotic plasmid is governed through the concerted action of one- and two-component signal transduction regulators. *Environ. Microbiol.* **15**, 811–821 (2013).
116. Heckel, B. C., Tomlinson, A. D., Morton, E. R., Choi, J.-H. & Fuqua, C. Agrobacterium tumefaciens exoR controls acid response genes and impacts exopolysaccharide synthesis, horizontal gene transfer, and virulence gene expression. *J. Bacteriol.* **196**, 3221–3233 (2014).
117. Ding, H. & Hynes, M. F. Plasmid transfer systems in the rhizobia. *Can. J. Microbiol.* **55**, 917–927 (2009).
118. Kumar, S. *et al.* N4-cytosine DNA methylation regulates transcription and pathogenesis in Helicobacter pylori. *Nucleic Acids Res.* **46**, 3429–3445 (2018).
119. Jones, P. A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492 (2012).
120. Gao, Q. *et al.* Bacterial DNA methyltransferase: A key to the epigenetic world with lessons learned from proteobacteria. *Front. Microbiol.* **14**, 1129437 (2023).

121. Sánchez-Romero, M. A. & Casadesús, J. The bacterial epigenome. *Nat. Rev. Microbiol.* **18**, 7–20 (2020).
122. Marinus, M. G. & Casadesus, J. Roles of DNA adenine methylation in host–pathogen interactions: mismatch repair, transcriptional regulation, and more. *FEMS Microbiol. Rev.* **33**, 488–503 (2009).
123. diCenzo George C. *et al.* DNA Methylation in Ensifer Species during Free-Living Growth and during Nitrogen-Fixing Symbiosis with Medicago spp. *mSystems* **7**, e01092-21 (2022).
124. Fioravanti, A. *et al.* DNA Binding of the Cell Cycle Transcriptional Regulator GcrA Depends on N6-Adenosine Methylation in Caulobacter crescentus and Other Alphaproteobacteria. *PLoS Genet.* **9**, (2013).
125. Pini, F. *et al.* Cell Cycle Control by the Master Regulator CtrA in Sinorhizobium meliloti. *PLOS Genet.* **11**, e1005232 (2015).
126. Shayanthan, A., Ordoñez, P. A. C. & Oresnik, I. J. The Role of Synthetic Microbial Communities (SynCom) in Sustainable Agriculture. *Front. Agron.* **4**, 896307 (2022).
127. Hacquard, S. & Schadt, C. W. Towards a holistic understanding of the beneficial interactions across the Populus microbiome. *New Phytol.* **205**, 1424–1430 (2015).
128. Hassani, M. A., Durán, P. & Hacquard, S. Microbial interactions within the plant holobiont. *Microbiome* **6**, 58 (2018).
129. Bardgett, R. D. & van der Putten, W. H. Belowground biodiversity and ecosystem functioning. *Nature* **515**, 505–511 (2014).

Appendix

Appendix – Phytoremediation: an interesting ability of plant microbiome to adapt to harsh environmental conditions

An expanding research area that covers air, soil and aquatic systems, astrobiology, biogeochemical cycles, plant-soil interactions, extreme environments, and spatial, temporal and perturbation studies is the environmental microbiology. It can be defined as the study of microbes, their functions, and interactions in all habitats on Earth [130]. Generally, microorganisms that inhabit unfavourable environments have the ability to rapidly evolve to adapt to harsh environmental conditions. Bacterial communities that live in natural extreme environments represent a gene reservoir for potential biotechnological applications Microorganisms adapted to extreme living conditions can provide research ideas for applications in bioremediation of polluted sites deemed too unbearable for most living organisms[131]. In the last couple of decades microbiome research has received tremendous attention and it has become evident that microbiota associated with higher organisms have highly important functions supporting health, growth, and well-being of their hosts[6].

Nevertheless, also human activities and climate changes have impacted terrestrial and marine ecosystems on global scale, resulting in new extreme environments. Despite their importance as key elements of ecosystem processes, microorganisms have been little studied in relation to the impact of human activities[132]. In particular, I studied how the plant microbiota reacts to external perturbations due to the use of these organisms in phytoremediation processes. The term phytodepuration refers to the plant-mediated remediation processes exploited to remove pollutants from water and wastewater and represents a promising technology for water reclamation because its economic and ecological sustainability. *Phragmites australis* is one of the most used plants. In recent years, the study of plant microbiota has received an increasing interest to understand the ability of plant to survive in hostile environments, such as that of constructed island (CW), which are engineered systems designed to mimic the natural processes of pollutants removal, maintaining a defined microbiota[133]. The perturbation of plant microbiota structure, or its stability, after the exposition to biotic and abiotic stresses is a promising research field to the exploitation of new bacterial strains to apply as biotechnological tool in phytoremediation processes. The use of specific bacterial inocula for microbial assisted phytodepuration has been proposed as a further advance for its implementation[132]. Nowadays, water reuse represents a priority to counteract water scarcity and minimize water footprint especially in agriculture, one of the highest water consuming human activities.

Following the aim to provide novel insights into microbiota of soils and associated with plants from "extreme" environments, I explored the bacterial communities of plants used in phytodepuration in the following paper.

xiv) Cangioli, L., Salobehaj, M., Del Duca, S., Fagorzi, C., Berardi, C., Coppini, E., Fibbi, D., Fani, R., Vassallo, A. (2022). Effect of Wastewater on the Composition of Bacterial Microbiota of Phragmites australis Used in Constructed Wetlands for Phytodepuration. Plants, 11 (23), art. no. 3210. doi: https://doi.org/10.3390/plants11233210

130. Watts, J. E. M. Welcome to Environmental Microbiome. *Environmental Microbiomes* Preprint at https://doi.org/10.1186/s40793-019-0340-8 (2019).
131. Mengoni, A., Schat, H. & Vangronsveld, J. Plants as extreme environments? Ni-resistant bacteria and Ni-hyperaccumulators of serpentine flora. *Plant Soil* **331**, 5–16 (2010).
132. Riva, V. *et al.* Microbial assisted phytodepuration for water reclamation: Environmental benefits and threats. *Chemosphere* **241**, 124843 (2020).
133. Vymazal, J. Constructed Wetlands for Wastewater Treatment: Five Decades of Experience. *Environ. Sci. Technol.* **45**, 61–69 (2011).

Overall activities

Publications

Number and order of publications refer to the text.

i. Cangioli, L., Mancini, M., Napoli, M., Fagorzi, C., Orlandini, S., Vaccaro, F., & Mengoni, A. (2022). Differential Response of Wheat Rhizosphere Bacterial Community to Plant Variety and Fertilization. International Journal of Molecular Sciences, 23(7), 3616. https://doi.org/10.3390/ijms23073616

ii. Cangioli, L., Mancini, M., Baldi, A., Fagorzi, C., Orlandini, S., Vaccaro, F., Mengoni, A. (2022). Effect of Site and Phenological Status on the Potato Bacterial Rhizomicrobiota. Microorganisms, 10 (9), art. no. 1743. https://doi.org/10.3390/microorganisms10091743

iii. Vaccaro F., Cangioli L., Mengoni A., Fagorzi C. (2022). Synthetic plant microbiota challenges in non-model species. Trends in Microbiology 30 (10), pp. 922-924. https://doi.org/10.1016/j.tim.2022.06.006

iv. Fagorzi, C., Passeri, I., Cangioli, L., Vaccaro, F., Mengoni, A. (2023). When biodiversity preservation meets biotechnology: The challenge of developing synthetic microbiota for resilient sustainable crop production. Journal of Sustainable Agriculture and Environment, 2 (1), pp. 5-15 https://doi.org/10.1002/sae2.12038

v. Bianco, C., Andreozzi, A., Romano, S., Fagorzi, C., Cangioli, L., Prieto, P., ... & Defez, R. (2021). Endophytes from African Rice (*Oryza glaberrima* L.) Efficiently Colonize Asian Rice (*Oryza sativa* L.) Stimulating the Activity of Its Antioxidant Enzymes and Increasing the Content of Nitrogen, Carbon, and Chlorophyll. Microorganisms, 9(8), 1714 https://doi.org/10.3390/microorganisms9081714

vi. Gullì M., Cangioli L., Frusciante S., Graziano S., Caldara M., Fiore A., Klonowski A.M., Maestri E., Brunori A., Mengoni A., Pihlanto A., Diretto G., Marmiroli N. and Bevivino A. (under submission). Combined metagenomic and metabolomic analyses to unlock the potential of co-applied SynComs and novel amendments from residual biomasses for a more sustainable agriculture.

vii. Fagorzi, C., Bacci, G., Huang, R., Cangioli, L., Checcucci, A., Fini, M., ... & Mengoni, A. (2021). Nonadditive transcriptomic signatures of genotype-by-genotype interactions during the initiation of plant-rhizobium symbiosis. Msystems, 6(1), e00974-20. https://doi.org/10.1128/msystems.00974-20

viii. Cangioli, L., Vaccaro, F., Fini, M., Mengoni, A., & Fagorzi, C. (2022). Scent of a Symbiont: The Personalized Genetic Relationships of Rhizobium—Plant Interaction. International Journal of Molecular Sciences, 23(6), 3358. https://doi.org/10.3390/ijms23063358

ix. Cangioli, L., Checcucci, A., Mengoni, A., & Fagorzi, C. (2021). Legume tasters: symbiotic rhizobia host preference and smart inoculant formulations. Biological Communications, 66(1), 47-54. https://doi.org/10.21638/spbu03.2021.106

x. Fagorzi, C., Ilie, A., Decorosi, F., Cangioli, L., Viti, C., Mengoni, A., & diCenzo, G. C. (2020). Symbiotic and Nonsymbiotic Members of the Genus *Ensifer* (syn. *Sinorhizobium*) Are Separated into Two Clades Based on Comparative Genomics and High-Throughput Phenotyping. Genome biology and evolution, 12(12), 2521-2534. https://doi.org/10.1093/gbe/evaa221

xi. diCenzo, G. C., Cangioli, L., Nicoud, Q., Cheng, J. H., Blow, M. J., Shapiro, N., ... & Mergaert, P. (2022). DNA methylation in *Ensifer* species during free-living growth and during nitrogen-fixing symbiosis with Medicago spp. Msystems, 7(1), e01092-21. https://doi.org/10.1128/mSystems.01092-21

xii. Riccardi, C., Passeri, I., Cangioli, L., Fagorzi, C., Fondi, M., Mengoni, A. (2023). Crossing Bacterial Genomic Features and Methylation Patterns with MeStudio: An Epigenomic Analysis Tool. International Journal of Molecular Sciences. https://doi.org/10.3390/ijms24010159

xiii. Passeri, I., Cangioli, L., Fondi, M., Mengoni, A., Fagorzi, C. (2023). The pan-epigenome of the symbiotic nitrogen fixing bacterium *Sinorhizobium meliloti* unravels unexpected variability of DNA-methylation profiles in closely related strains. bioRxiv 2023.05.31.542881; doi: https://doi.org/10.1101/2023.05.31.542881

xiv. Cangioli, L., Salobehaj, M., Del Duca, S., Fagorzi, C., Berardi, C., Coppini, E., Fibbi, D., Fani, R., Vassallo, A. (2022). Effect of Wastewater on the Composition of Bacterial Microbiota of *Phragmites australis* Used in Constructed Wetlands for Phytodepuration. Plants, 11 (23), art. no. 3210. doi: https://doi.org/10.3390/plants11233210

Other publications:

- Bacci, G., Rossi, A., Armanini, F., Cangioli, L., De Fino, I., Segata, N., ... & Bevivino, A. (2021). Lung and Gut Microbiota Changes Associated with *Pseudomonas aeruginosa* Infection in Mouse Models of Cystic Fibrosis. International journal of molecular sciences, 22(22), 12169.

Work in progress

I. Synergistic effect of rhizobia and plant-growth promoting nonsymbiotic bacteria on alfalfa growth

II. Study of the regulatory cascade involved in the megaplasmids conjugation in the model plant symbiotic bacterium *Sinorhizobium meliloti* by the creation of a GFP-based reporter system

III. The impact of synthetic microbial consortia, fertilization regimes and maize growth stages on taxonomic and functional microbiome under temperate climate conditions

IV. Genomic characterisation of bacterial strains composing an effective synthetic community applied as biofertilizer for maize, wheat and tomato

Conferences

1. A. Bevivino, J. Hett, L. Cangioli, A. Fiore, N. Kleinbölting, A. Mengoni , M. Costanzo, A. Schlueter, D Neuhoff, A. Sczyrba. The impact of synthetic microbial consortia, fertilization regimes and maize growth stages on plant growth and rhizosphere microbiome under temperate climate conditions *Continental IUSS 2024 (19th-21st May 2024), Florence, Italy*

2. L. Cangioli, M. Fini, G. Cambi, A. Mengoni, C. Fagorzi Poster session Study of the regulatory cascade involved in the megaplasmids conjugation in the model plant symbiotic bacterium Sinorhizobium meliloti by the creation of a GFP-based reporter system *XXXIV SIMGBM Congress (21st – 24th September 2023) MICROBIOLOGY 2023, Cagliari, Italy*

3. A. Mengoni, L. Cangioli, I. Passeri, F. Vaccaro, C. Fagorzi When biodiversity preservation meets biotechnology: the challenge of developing synthetic microbiota and the case of symbiotic nitrogen-fixing bacteria *Workshop "Trends in microbial solutions for sustainable agriculture" 14 – 15 September 2023, Belgrade, Serbia*

4. M.L. Amenta, S. Varriale, L. Cangioli, C. Fagorzi, G. Amato, D. Giambalvo, R. Defez, A. Mengoni, C. Bianco Cereal-associated bacterial endophytes promote nitrogen fixation in rhizobium-legume symbiosis *15th European Nitrogen Fixation Conference (31st August – 3rd September 2023), Naples, Italy*

5. A. Bevivino, J. Hett, L. Cangioli, A. Fiore, N. Marmiroli, D Neuhoff. Unlocking the potential of microbiome-based solutions as green biofertilizer for sustainable agriculture *FEMS 2023 (9th-13th July 2023), Hamburg, Germany*

6. L. Cangioli, M. Fini, G. Cambi, A. Mengoni , C. Fagorzi Study of the regulatory cascade involved in the megaplasmids conjugation in the model plant symbiotic bacterium *Sinorhizobium meliloti* by the creation of a GFP-based reporter system *FISV 2022 (14th – 16th September 2022), Naples, Italy*

7. Antonia Esposito, L. Cangioli, A. Vassallo, S. Del Duca, G. Semenzato, C. Berardi, E. Coppini, D. Fibbi, R. Fani Temporal evolution of bacterial communities associated to Phragmites australis exploited in phytodepuration of wastewater *FISV 2022 (14th – 16th September 2022), Naples, Italy*

8. C. Fagorzi, I. Passeri, C. Riccardi, L. Cangioli, M. Fondi, A. Mengoni
   Unravelling the bacterial epigenome and its influence on gene transfer and transcriptional control *FISV 2022 (14th – 16th September 2022), Naples, Italy*

9. L. Cangioli, M. Mancini, M. Napoli, C. Fagorzi, S. Orlandini, F. Vaccaro, A. Mengoni
   Differential response of wheat rhizosphere microbiotabacterial community to plant variety and fertilization *miCROPe Symposium 2022 (11th – 14th July 2022), Wien, Austria*

10. Bevivino, A. Fiore, J. Hett, M. Caldara, L. Cangioli, E. Maestri, G. Aprea, M. Costanzo, M. Gullì, S. Graziano, A. Mengoni, S. Pihlanto, D. Neuhoff, N. Marmiroli Effect of Plant Probiotics and Biochar Application on Rhizosphere Biodiversity and Maize Growth in the Field miCROPe Symposium 2022 (11th – 14th July 2022), Wien, Austria

11. L. Cangioli, A. Fiore, J. Hett, M. Caldara, E. Maestri, G. Aprea, M. Costanzo, M. Gullì, S. Graziano, A. Mengoni, A. Pihlanto, D. Neuhoff, N. Marmiroli and A. Bevivino Effect of Plant Probiotics and Biochar Application on Rhizosphere Biodiversity and Maize Growth in the Field *CortonaProcarioti2022 (23rd – 25th June 2022), Cortona, Italy*

12. L. Cangioli, A. Fiore, J. Hett, M. Caldara, E. Maestri, G. Aprea, M. Costanzo, M. Gullì, S. Graziano, A. Mengoni, A. Pihlanto, D. Neuhoff, N. Marmiroli A. Bevivino
    Effect of Plant Probiotics and Biochar Application on Rhizosphere Biodiversity and Maize Growth in the Field
    CortonaProcarioti2022 (23rd – 25th June 2022), Cortona, Italy

13. Bevivino, D. Neuhoff, N. Marmiroli, J. Hett, A. Fiore, G. Aprea, A. Mengoni, L. Cangioli, M. Caldara, E. Maestri, P. Ambrosino, S. Passato, S. Frusciante, S. Tabacchioni, A. Del Fiore, A. Ambrico, A. Magarelli, M. Trupo, M. Costanzo, M. Gullì, S. Graziano, G. Giovannetti, G. Masoero, A. Brunori, A. Pihlanto.
    Application of Beneficial Microbial Consortia as Plant Probiotics for a Sustainable Agriculture
    Microbial Diversity 2021 Congress (14th – 15th December 2021), virtual conference

14. G. diCenzo, L. Cangioli, Q. Nicoud, J. Cheng, M. Blow, N. Shapiro, T. Woyke, B. Alunni, A. Mengoni, P. Mergaert The DNA methylome of Sinorhizobium meliloti and its changes during symbiosis with legumes
    14th European Nitrogen Fixation Conference (22nd September – 22nd October 2021), virtual conference

15. Fagorzi, A. Bellabarba, L. Cangioli, G. diCenzo, G. Bacci, C. Viti, F. Pini, A. Mengoni
    Legume tasters: The panoply of symbiotic rhizobia host preference *14th European Nitrogen Fixation Conference (22nd September – 22nd October 2021), virtual conference*

16. Fagorzi, G. Bacci, R. Huang, L. Cangioli, A. Checcucci, G. diCenzo, A. Mengoni
    On the search of the legume tasters: Genotype-by-Genotype Interactions during the Initiation of Plant-Rhizobium Symbiosis *14th European Nitrogen Fixation Conference (22nd September – 22nd October 2021), virtual conference*

17. Fagorzi, I. Passeri, C. Riccardi, L. Cangioli, M. Fondi, G. diCenzo, A. Mengoni
    Unraveling the bacterial epigenome and its influence on gene transfer and transcriptional control *SMRT Leiden 2021 – Young Investigator Virtual Conference (26th – 27th May 2021),* virtual conference

Courses

- *Introduzione agli strumenti per una Smart agriculture - AGRISMART"*
  Edizione Università di Firenze (23rd - 24th February 2021)
- *SIMBA / de.NBI Metagenomics Workshop*
  Online training course 3rd - 5th February 2021

International Journal of
Molecular Sciences

MDPI

Article

Differential Response of Wheat Rhizosphere Bacterial Community to Plant Variety and Fertilization

Lisa Cangioli [1], Marco Mancini [2], Marco Napoli [2,*], Camilla Fagorzi [1], Simone Orlandini [2], Francesca Vaccaro [1] and Alessio Mengoni [1,*]

[1] Department of Biology, University of Florence, Via Madonna del Piano 6, 50019 Sesto Fiorentino, Italy; lisa.cangioli@unifi.it (L.C.); camilla.fagorzi@unifi.it (C.F.); francesca.vaccaro@unifi.it (F.V.)
[2] Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, Piazzale delle Cascine 18, 50144 Florence, Italy; marco.mancini@unifi.it (M.M.); simone.orlandini@unifi.it (S.O.)
* Correspondence: marco.napoli@unifi.it (M.N.); alessio.mengoni@unifi.it (A.M.)

Abstract: The taxonomic assemblage and functions of the plant bacterial community are strongly influenced by soil and host plant genotype. Crop breeding, especially after the massive use of nitrogen fertilizers which led to varieties responding better to nitrogen fertilization, has implicitly modified the ability of the plant root to recruit an effective bacterial community. Among the priorities for harnessing the plant bacterial community, plant genotype-by-microbiome interactions are stirring attention. Here, we analyzed the effect of plant variety and fertilization on the rhizosphere bacterial community. In particular, we clarified the presence in the bacterial community of a varietal effect of N and P fertilization treatment. 16S rRNA gene amplicon sequence analysis of rhizospheric soil, collected from four wheat varieties grown under four N-P fertilization regimes, and quantification of functional bacterial genes involved in the nitrogen cycle (*nifH*; *amoA*; *nirK* and *nosZ*) were performed. Results showed that variety played the most important role and that treatments did not affect either bacterial community diversity or bacterial phyla abundance. Variety-specific response of rhizosphere bacterial community was detected, both in relation to taxa (Nitrospira) and metabolic functions. In particular, the changes related to amino acid and aerobic metabolism and abundance of genes involved in the nitrogen cycle (*amoA* and *nosZ*), suggested that plant variety may lead to functional changes in the cycling of the plant-assimilable nitrogen.

Keywords: wheat; rhizosphere; bacterial community; ecosystem functions

check for updates

Citation: Cangioli, L.; Mancini, M.; Napoli, M.; Fagorzi, C.; Orlandini, S.; Vaccaro, F.; Mengoni, A. Differential Response of Wheat Rhizosphere Bacterial Community to Plant Variety and Fertilization. *Int. J. Mol. Sci.* **2022**, *23*, 3616. https://doi.org/10.3390/ijms23073616

Academic Editors: Yangrong Cao and Hari B. Krishnan

Received: 10 February 2022
Accepted: 23 March 2022
Published: 25 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The plant bacterial community, and especially the rhizospheric bacterial community, has been clearly related to crop yield and resistance to biotic and abiotic stresses [1–6]. Domestication and breeding have strongly influenced plant bacterial community [7,8], resulting in some cases in a reduced ability of the plant to attract an adequate rhizosphere bacterial community [9]. Since the diversity of the plant bacterial community is contributing to the resilience and sustainability of agroecosystems to multiple stressors, including climate change, increased diversity of the crop-associated bacterial community can contribute to conjugating sustainable yield with a reduced footprint of agriculture on the environment. Indeed, in general terms, plant genotype is a relevant factor affecting diversity and functions of the plant bacterial community and influencing soil bacterial community [3,10–12]. Consequently, there is a growing focus on considering the recruited plant bacterial community for future crop breeding programs [3]. Among the priorities for harnessing the plant bacterial community, plant genotype × microbiome × environment × management interactions are stirring attention [13]. In fact, not only the species, but also the individuality of genotypes, viz. germplasm or varieties, show peculiarities in the associated bacterial community stresses [5,8,14].

With the increasing food demand [15,16], winter wheat grain production needs to expand by 11% by 2026 [17]. To increase the grain yield and protein production per hectare, intensive breeding of winter wheat started in the second half of the twentieth century [18]. Winter wheat cultivars before the late 1960s are commonly called "old", while those registered later are called "modern" [19]. In particular, the introduction of growth repressor (Rht) genes led to the development of semi-dwarf genotypes [20] which are characterized by improved efficiency of partitioning photosynthetic assimilates and lodging resistance [18,21], are highly productive, homogeneous, with high protein production per hectare, and with a gluten composition suitable for industrial processing [19,22–24]. On the other hand, the intensive breeding of winter wheat led to lower vitamin and micronutrient concentrations in grains [25] as a consequence of the "yield dilution phenomenon" in "modern" varieties [26]. Furthermore, the smaller root system induced by the dwarfism gene [27] makes the "modern" varieties less able to explore the soil and therefore more reliant on chemical fertilization than the "old" ones [28]. Further, the breeding process may also have influenced the characteristics and quantity of root exudates in "modern" varieties, thus influencing the composition of the rhizosphere [20]. These features suggest that "old" and "modern" wheat varieties may harbor different root-associated microbiota related to the different nutrient requirements and root systems. In fact, concerning wheat-associated bacterial communities, a clear effect of variety on the rhizosphere bacterial community has been shown [29–31], as well as nitrogen fertilization [20,32]. Recently, a hypothesis of reduction of the level of plant-bacterial community interaction along wheat domestication has been proposed [33,34]. In this context, bacterial communities in the rhizosphere of "old" wheat varieties were found to be more taxonomically diverse than that detected for "modern wheat", probably due to differences in root exudates [35].

Given the resurged importance of old wheat varieties as a source of germplasm for breeding to develop varieties with higher nutritional value and better use of soil nutrients, here, we wanted to identify in a panel of three old and one modern variety if the wheat rhizosphere bacterial community was more affected by the variety or by the fertilization. In particular, we aimed to investigate the presence of a differential effect that the same N and P fertilization treatment may have on different varieties. We answered this question by performing 16S rRNA gene amplicon sequencing analysis of rhizospheric soil, collected from three old and one modern wheat varieties grown under four N-P fertilization regimes. Taxonomic profiling and putative functional attribution were produced, as well as quantification of functional bacterial genes involved in the nitrogen cycle (nitrogen fixation, *nifH*; nitrification, *amoA*; denitrification, *nirK* and *nosZ*), to estimate the combined effect of wheat variety and fertilization on rhizosphere bacterial community ecosystem functions.

2. Results

2.1. Physicochemical Characterization of the Soil

Results of the physicochemical analysis of the soil samples are reported in Table 1. For each plot, the analysis was performed on bulk soil samples. A Principal Component Analysis (PCA) was performed on soil physicochemical parameters and the location of sample collection (i.e., latitude and longitude of the plots). No evident clustering was detected, either with respect to physicochemical parameters and geographical coordinates (Figure 1).

Table 1. Results of the soil physicochemical analysis. Sample codes refer to the name of the samples reported in Table 1. For each soil parcel analysis was performed on bulk soil ("S") samples. Exc, exchangeable.

	Sample Codes															
	S5S	**V2S**	**B14S**	**A11S**	**A12S**	**B13S**	**V1S**	**S6S**	**S7S**	**S8S**	**V4S**	**V3S**	**B16S**	**B15S**	**A9S**	**A10S**
Organic C (g/kg)	0.9	0.7	0.8	1	1.1	1	0.8	0.9	0.8	0.8	1	0.7	0.9	1	1	1.1
Organic matter (g/kg)	1.6	1.3	1.4	1.7	1.8	1.7	1.4	1.6	1.4	1.4	1.7	1.2	1.5	1.7	1.8	1.9
Total N (g/kg)	1.1	1	1.1	1.1	1.2	1.2	1	1.1	1.3	1.1	1.2	1	1	1.2	1.1	1.2

Table 1. *Cont.*

	Sample Codes															
	S5S	**V2S**	**B14S**	**A11S**	**A12S**	**B13S**	**V1S**	**S6S**	**S7S**	**S8S**	**V4S**	**V3S**	**B16S**	**B15S**	**A9S**	**A10S**
Nitric N (mg/kg)	4	4	4	4	4	4	4	4	4	4	4	4	6	4	4	4
Ammoniacal N (mg/kg)	6	2	4	3	6.4	5	5	6.5	4	4	5	3	5.6	6	4	4
Assimilable P (mg/kg)	5	4	4	5	6	4	4	9	4	4	4	11	7	4	5	6
Exc. P (as K_2O) (mg/kg)	569	316	321	381	407	378	313	487	378	376	607	362	559	436	591	607
Exc. P (mg/kg)	236	131	133	158	169	157	130	202	157	156	252	150	232	181	245	252
Total limestone (g/kg)	12	15	15	12	9	13	12	13	12	11	11	15	18	13	14	15
ExcCa (mg/kg)	3497	2768	2746	3279	3538	2798	2640	2962	2778	2594	3140	3111	2995	2918	3095	3116
Exc. Ca (as CaO) (mg/kg)	4895.8	3875.2	3844.4	4590.6	4953.2	3917.2	3696	4146.8	3889.2	3631.6	4396	4355.4	4193	4085.2	4333	4362.4
Exc. Mg (mg/kg)	153	96	84	119	114	106	88	112	104	94	147	75	155	134	163	149
Exc. Mg (as MgO) (mg/kg)	255	160	140	198	190	177	147	187	173	157	245	125	258	223	272	248
Exc. Na (mg/kg)	18	13	16	13	17	15	15	19	19	19	18	16	20	17	21	19
Assimilable Fe (mg/kg)	12.5	11	11	11	9.1	12	12	9.3	11	10	12	10	8.5	10	11	11
Assimilable Zn (mg/kg)	0.4	0.3	0.4	0.4	0.3	0.4	0.3	0.6	0.5	0.4	0.5	1	0.5	0.6	0.5	0.8
Assimilable Se (mg/kg)	0.3	0.2	0.1	0.2	0.1	0.1	0.1	0.1	0.2	0.3	0.1	0.1	0.2	0.1	0.1	0.1
pH in water	8.2	8.2	8.2	8.2	8.2	8.1	8	8.2	8.2	8.2	8.1	8.2	8.2	8.1	8.3	8.3
C.E.C. pH 8.2 (meq/100 g)	19	15	15	18	19	15	14	16	15	14	18	17	17	16	18	18

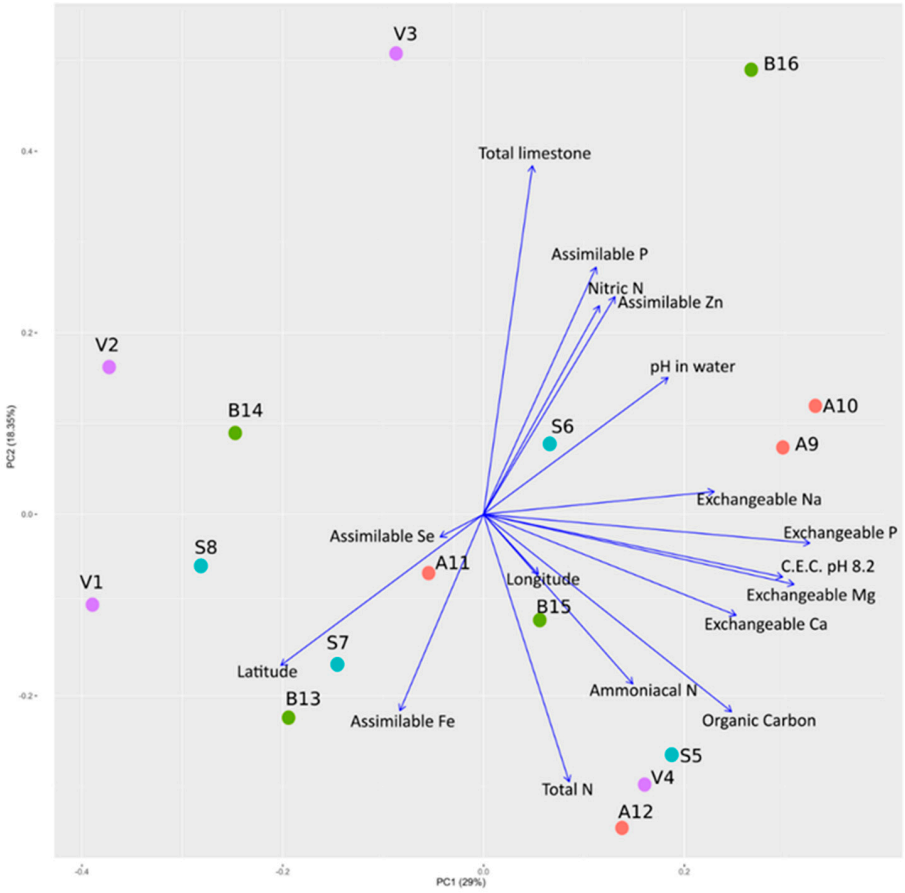


Figure 1. Influence of physicochemical composition and geographical coordinates of soil samples with respect to the sowed varieties. A Principal Component Analysis of physicochemical composition and geographical coordinates of soil samples is reported with a biplot showing the percentage of variance of the first two components. Codes of samples are as in Table 1.

2.2. Diversity of the Soil and Rhizosphere Bacterial Community

A total of 3,180,595 16S rRNA reads were obtained, and 2,131,486 (67% of total reads) passed quality filtering (Table S1). After the clustering step, a total of 12,795 amplicon sequence variants (ASVs) were obtained (Dataset S1). To Bacteria, 10,271 ASVs were

assigned, 59 ASVs to Archaea, and 2,465 were unassigned (NA). ASVs assigned to Archaea and NA were removed from the following analyses. Rarefaction curves obtained from ASVs reached plateaus for all samples (Figure S1), indicating a satisfactory survey of the bacterial diversity (Good's coverage Table S2), which allowed to estimate alpha diversity indices (Table S3). Bulk soil and rhizosphere samples did not differ in alpha diversity nor in relation to both variety and fertilization (Permanova $p > 0.05$). Varieties differed in relation to the Shannon index, but not in relation to the Simpson index (Figure 2, see specific pairwise comparison for *p*-values), indicating that species richness and species abundance distribution (equitability) changed among varieties, while the dominance of bacterial groups did not.

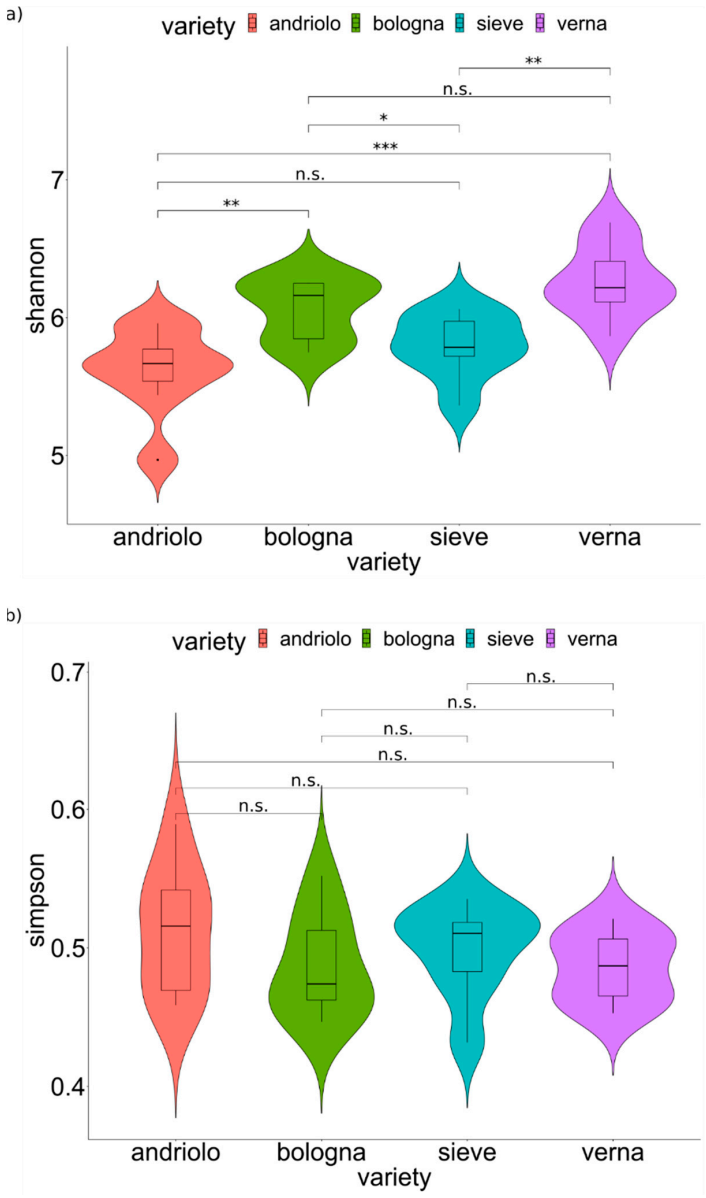


Figure 2. Diversity of soil and rhizosphere bacterial community. Values for alpha diversity indices are reported. Violin plots with samples (bulk and rhizosphere soil) are grouped according to varieties. Statistical analyses were performed using the Wilcoxon test (*p*-values are reported over the lines: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; n.s., not significant $p > 0.05$). The y axis indicates the values of the Shannon index (**a**) and the values of the Simpson index (**b**).

Concerning the taxonomic diversity among samples, alpha and beta diversities of soil and rhizosphere samples were not separated (Figure 3), while plant varieties were, with evident grouping for Andriolo and Bologna, and less evident, but significant for Verna and Sieve (Table S4).

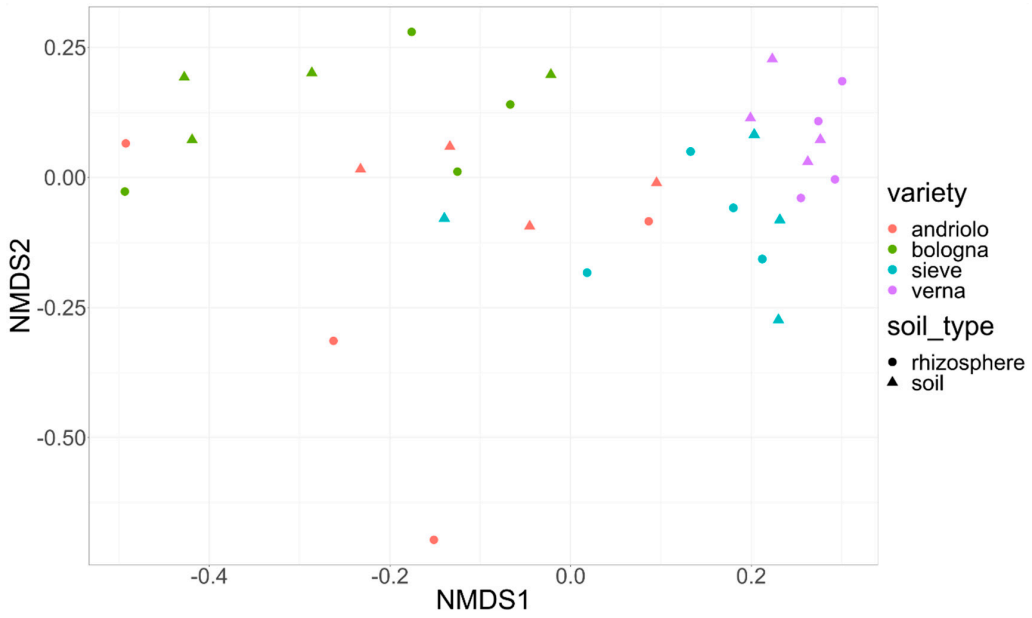


Figure 3. Soil sample distribution according to bacterial community. Nonmetric Multidimensional Scaling of soil and rhizosphere bacterial community. Ordination based on Bray–Curtis distance. Colors indicate the wheat variety, and shapes refer to the type of sample (circle = rhizosphere soil, triangle = bulk soil). Stress = 0.15.

Fertilization did not affect the taxonomic diversity of the samples. Statical analysis was performed using a permutational multivariate analysis of variance (PERMANOVA) on centered log-ratio-transformed samples (Table S4).

2.3. Variety Affects the Taxonomic Composition of Wheat Rhizobacterial Community

Figure 4 reports the relative abundances of taxonomies at the phylum level for the top 20 ASVs detected (which represent relative abundances from 0.1% to 0.08%). Taxonomic abundances in relation to variety and fertilization are shown in Figure S2. The most represented phyla were Acidobacteriota, Actinobacteriota, Chloroflexi, Proteobacteria, and Verrucomicrobia. Differences among varieties were present in relation to phyla abundance (Figure 5). In fact, differential abundance analyses showed remarkable differences in Actinobacteriota, Chloroflexi, Desulfobacterota, Entotheonellaeota, Firmicutes, Myxococcota, Nitrospirota, and the candidate division RCP2-54, in particular for Verna variety, as confirmed by a Permanova on the PCA to the centered log-ratio (Figure S3). When considering single ASVs, 37 of them were found with statistically different abundance in relation to variety (Table S5). These ASVs included members of the orders Burkholderiales, Rhizobiales (Beijerinckiaceae, Hyphomicrobiaceae, Xanthobacteraceae), Blastocatellales, Pyrinomonadales, as well as of Actinobacteriota.

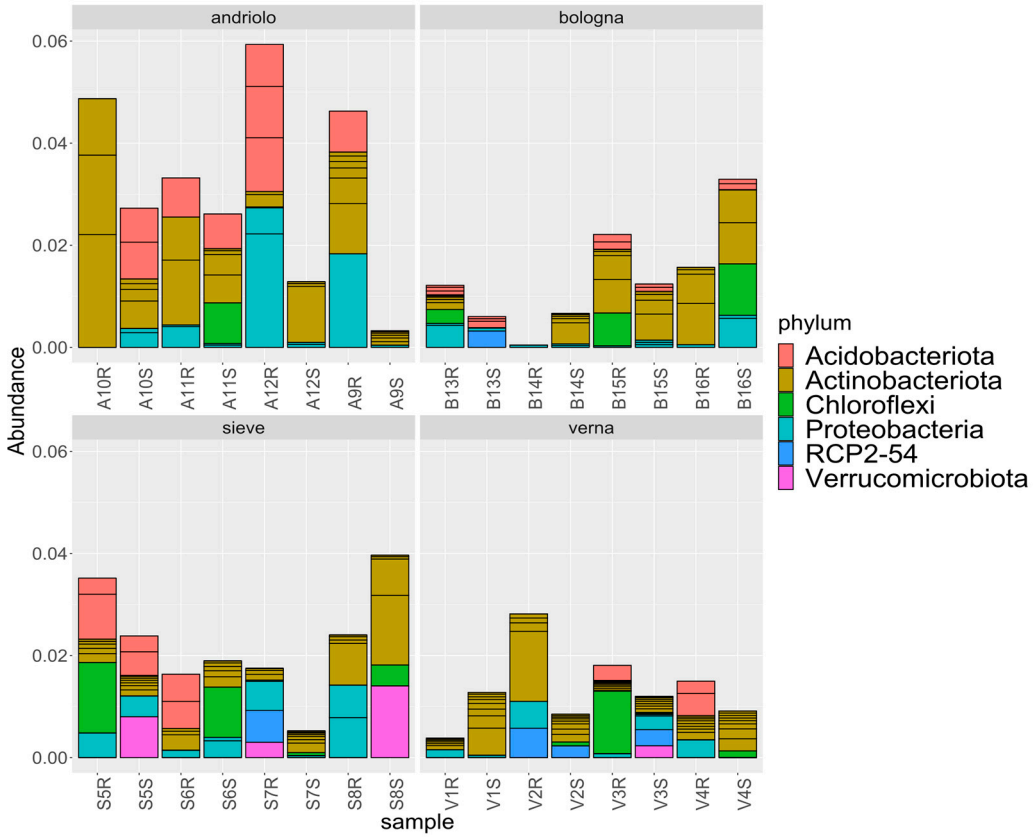


Figure 4. Relative abundances of taxonomies. Abundances of the top twenty detected ASVs are reported at phylum level in relation to variety.

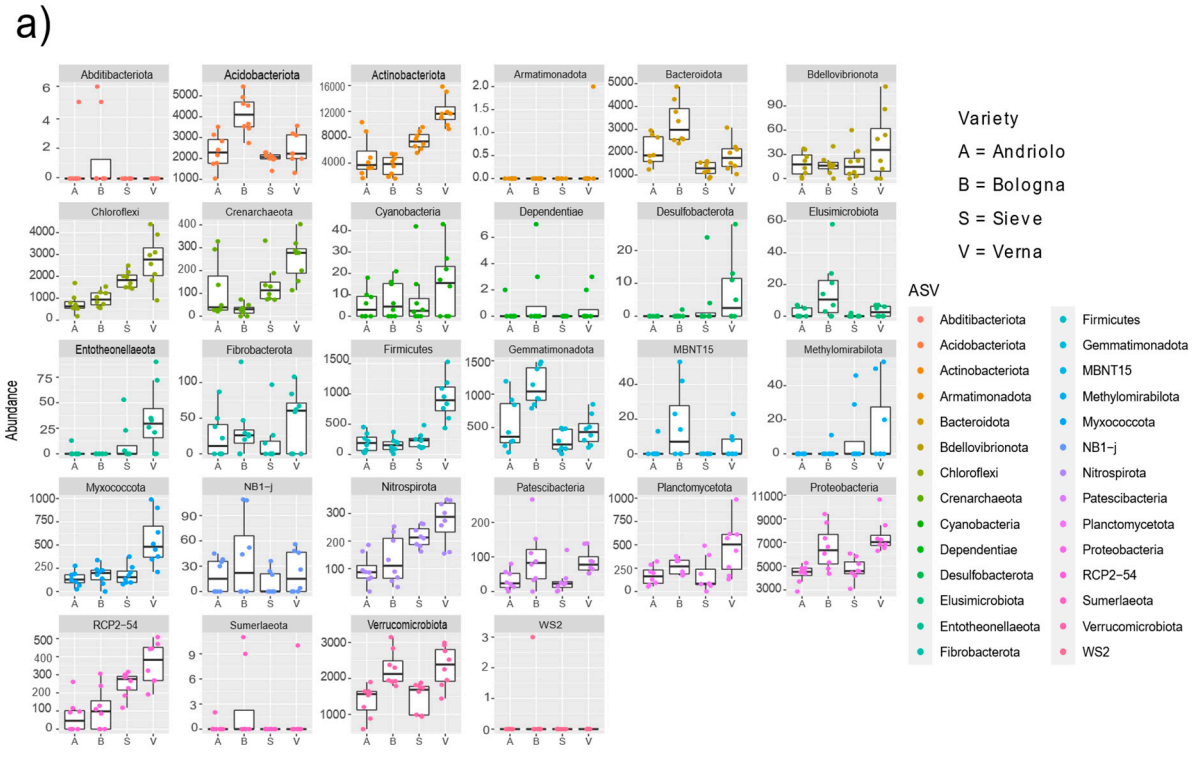


Figure 5. *Cont.*

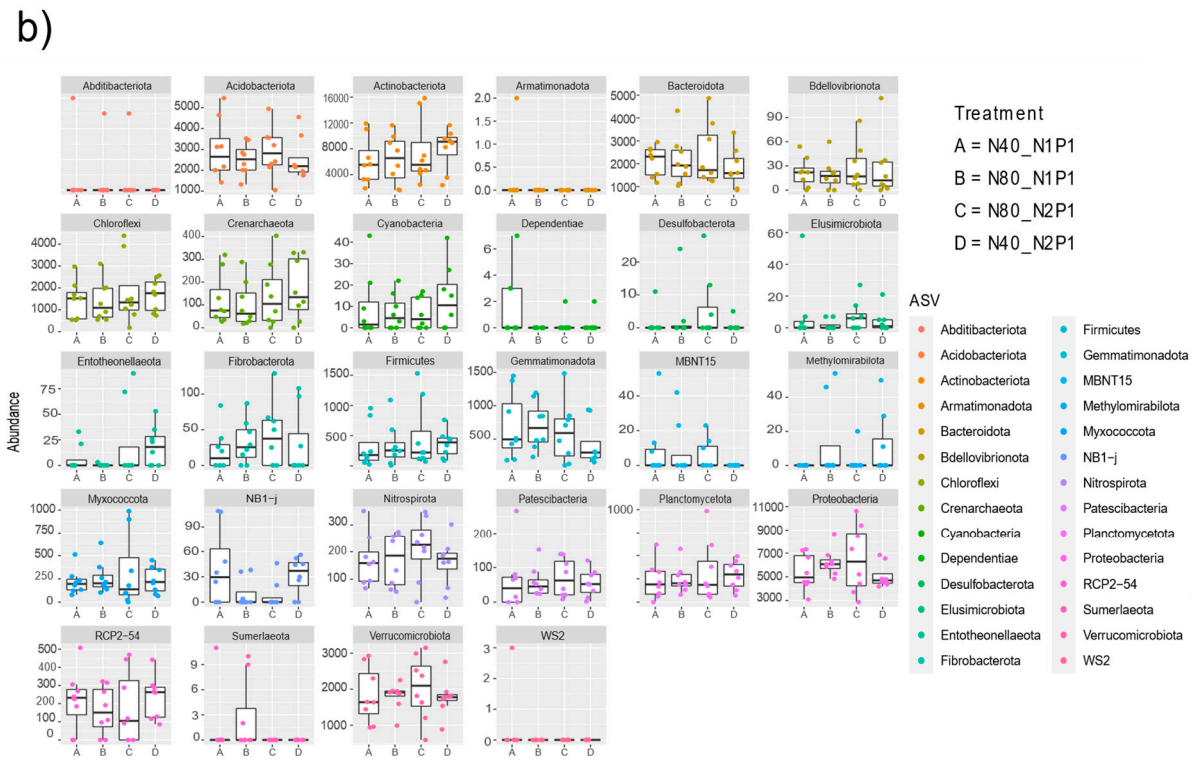


Figure 5. Variety and fertilization affect specific bacterial phyla abundances. Differential abundances of bacterial phyla in relation to wheat varieties and fertilization is reported from DeSeq2 analysis result. For each phylum, the abundance is reported on the y axis with respect to the varieties (**a**) and the fertilization (**b**).

2.4. Prediction of Potential Functions and Metabolic Pathways

Functional profiles of the bacterial community were inferred. All the identified metabolic pathways present in the taxa retrieved in the bacterial community are listed in Table S6. We identified a total of 422 pathways, all distributed in all samples. However, pathways' abundance showed a clustering according to varieties (Figure 6a). Using a SIMPER test, we then selected the top three pathways that mostly differentiate the four varieties (in terms of amount of variance explained). Six different pathways-(VALSYN-PWY, ILEUSYN-PWY, BRANCHED-CHAIN-AA-SYN-PWY, PWY-3001, PWY-3781, and PWY-7111) were selected, most of them were related to the metabolism of amino acids and aerobic respiration. In particular, the pathways of L-valine biosynthesis (VALSYN-PWY), L-isoleucine biosynthesis (ILEUSYN-PWY and PWY-3001) and aerobic respiration (PWY-3781) were found. Interestingly, this subset of pathways also suggested a clustering related to, in particular, Verna being grouped apart from the other three varieties (Figure 6b). Pathways related to the nitrogen cycle showed only a minor contribution to the overall differences (i.e., the denitrification pathway accounted to 0.05% of total variance).

2.5. Quantitative PCR Analysis of the Nitrogen Cycle Genes

In order to overcome the limitation of PICRUSt-based pathway predictions (which are based on overall pathways and on phylogenetic reconstruction) and more clearly define the effect of plant variety and fertilization on the bacterial community fraction involved in the nitrogen cycle, bacterial genes related to nitrogen fixation (from N_2 to NH_3, *nifH*), nitrification (conversion of ammonia $NH_3$to hydroxylamine, *amoA*), and denitrification (nitrite to nitric oxide, *nirK* and nitrous oxide to N_2, *nosZ*) were detected, and their abundance was estimated by quantitative PCR on rhizosphere samples (Figure 7). While for *nifH* no differences among varieties or fertilizations were detected, *amoA*, *nirK*, and *nosZ* were different with respect to varieties (Scott–Knott test $p < 0.05$, Table S7). In particular,

Andriolo and Bologna showed higher abundance of *nirK* and *amoA* genes than Sieve and Verna did. Verna also had lower *nosZ* values than the other three varieties did, suggesting that the denitrification pathway could be less effective on the rhizosphere of this variety.

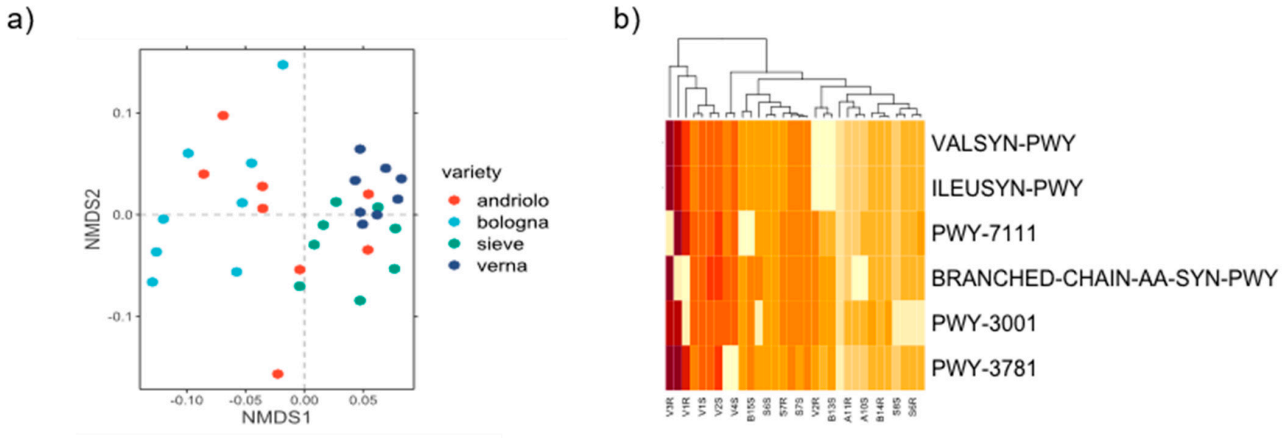


Figure 6. Clustering of pathways abundance. In panel (**a**), Nonmetric Multidimensional Scaling analysis (nMDS) was performed considering the predicted functional abundances of all identified pathways. In panel (**b**), a heatmap considering the top three pathways that differentiated each contrast between varieties returned by SIMPER test is represented. Codes of represented pathways correspond to the following pathway descriptions: L-valine biosynthesis (VALSYN-PWY), L-isoleucine biosynthesis I (from threonine) (ILEUSYN-PWY), pyruvate fermentation to isobutanol (PWY-7111), superpathway of L-leucine, L-valine, and L-isoleucine biosynthesis (BRANCHED-CHAIN-AA-SYN-PWY), L-isoleucine biosynthesis III (PWY-3001) and aerobic respiration I (cytochrome c) (PWY-3781).

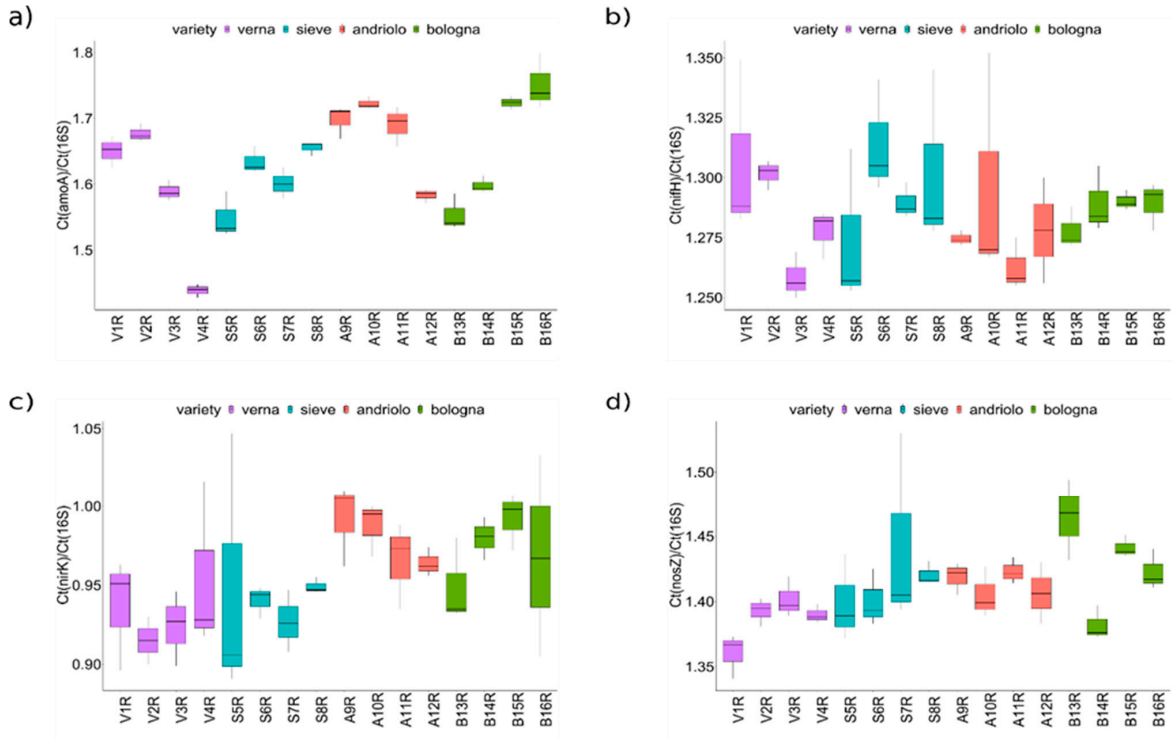


Figure 7. Abundance of genes involved in the nitrogen cycle. Plots report the averages of the relative gene quantification with respect to the 16S rRNA gene. Ratios of gene copies are a proxy of the abundances of bacteria harboring the indicated genes with respect to the total number of bacteria: (**a**) *amoA* gene; (**b**) *nifH* gene, (**c**) *nirK* gene, and (**d**) *nosZ* gene.

3. Discussion

The taxonomic assemblage and functions of the plant bacterial community are strongly influenced by soil and host plant genotype [5,7,8]. Crop breeding, especially after the massive use of nitrogen fertilizers which led to varieties responding better to nitrogen fertilization [36], has implicitly decreased the pressure toward plant genotypes highly efficient in the use of native soil nutritional resources, then indirectly reducing the ability of the plant root to recruit an effective bacterial community [7,8,37]. Tough physicochemical analysis of soil did not detect evident groupings according to wheat varieties (apart from Andriolo), and the analysis of the bacterial community did, clearly differentiating the four varieties under study (Andriolo being the most heterogeneous, Figure 3). In fact, the overall body of our results showed that variety played the most important role and that fertilization did not affect either bacterial community diversity or bacterial phyla abundance. Interestingly, rhizosphere bacterial community and bulk soil bacterial community were highly similar, suggesting that under the plant density and the depth we sampled, the root apparatus could extend its effects at a certain distance from the plant stem (30 cm). This bacterial community was in agreement with the extensive root apparatus of wheat, which allows affecting the soil at a distance from the plant [38]. Consequently, we may consider "bona fide" our bulk soil samples as very close to a low-bound rhizosphere soil. It would be interesting to evaluate the extent in space around the plant of this assimilation of soil on the rhizosphere bacterial community, which can have relevance for the setting-up and interpretation of results from intercropping trials [39]. The three old varieties and the modern Bologna variety showed a differential abundance of several microbial taxa. Among them, Acidobacteriota, Bacteroidota, Chloroflexi, Firmicutes, and Nitrospirota were particularly relevant. Previous studies on wheat rhizosphere showed that N fertilization decreased the abundance of Acidobacteria and increased Bacteroidetes [20]. In our work, a decrease of Acidobacteriota was shown by Andriolo variety in the plots with increased P-fertilization (Figure S2), which may suggest that this variety may respond differently to P fertilization, in turn affecting its rhizosphere bacterial community. However, both Andriolo and Bologna showed levels of abundance of members of phyla Acidobacteriota and Bacteroidota higher than that of Sieve and Verna varieties. Since both Sieve and Verna derive from breeds with "Est Mottin 72", we could not a priori exclude that the lower levels of Acidobacteriota and Bacteroidota in their rhizosphere may reflect this partially common genetic basis, highlighting a contribution of plant genotype to the interaction with specific bacterial phyla. If confirmed in following experiments, these bacterial phyla could be good candidates for quantitative trait loci (QTL) analysis identifying in breeds from the old Verna variety and the modern Bologna. The high abundance of members of Acidobacteriota also agrees with previous reports (see, for instance, [33,40]). Moreover, Bacteroidota, Chloroflexi, and Firmicutes phyla were recognized as important in differentiating wheat species (*T. durum* Desf. and *T. aestivum* L.) [33] and soil-varieties combinations [31], suggesting the presence of loci in the wheat genome which can be identified for interaction with the rhizosphere microbiota. To date, few loci/genes in a relatively small number of plant species have been identified as related to the recruitment of microbiota [41], and no studies have been performed on wheat. Interestingly, ASVs affiliated to the uncultured phylum RCP2-54 were found more abundantly in Sieve and mainly in Verna, compared to Bologna and Andriolo. The functional role of this phylum is still unknown; however, previous works showed its presence in the rhizosphere and in differentiating tomato and maize varieties under water deficit conditions [42], again suggesting a possible genetic basis in the plant for its differential recruitment. At the level of ASVs it was relevant to find members of Blastocatellales and Pyrinomonadales. A recent paper investigating soil suppressiveness to *Fusarium culmorum* in wheat, found Blastocatellales in the network associated with pathogen suppression [43], and Pyrinomonadales were reported to vary in abundance in wheat rhizosphere in relation to soil nitrate level [44]. This evidence may suggest that the four varieties in our study may be able to differentially recruit a fraction of the soil bacterial community able to improve plant health and contributing differently to plant

use of nitrogen. Bologna variety had the highest levels of both Blastocatellales and Pyrinomonadales, although high variation between plots was present (Figure S4). The role of these bacterial groups on plant tolerance to pathogenic fungi and nitrogen use may deserve further attention.

Under the light of possible functional differences among the bacterial community recruited by the four plant varieties, PICRUSt2 results (Figure 6) mirrored in part the taxonomic profiling in varieties differentiation. In particular, Verna grouped separately from the other three varieties, similarly to the pattern shown in the nMDS analysis of ASVs (Figure 3). Interestingly, the metabolic pathways which mostly differentiated Verna from the other varieties were related to aerobic respiration and amino acid biosynthesis. We may speculate that root exudates from the four varieties differed in the amino acid composition, hence differentially promoting the growth of bacteria harboring specific biosynthesis pathways. Indeed, root exudates are rich in amino acids, and previous work showed that amino acid content strongly varies among different plant varieties. Concerning the aerobic respiration pathway, Verna showed a more abundant representation than the other varieties did, though with differences among fertilization regimes. This may suggest that the other varieties could hold a higher fraction of anaerobic bacteria/anaerobic metabolism. Verna rhizosphere bacterial community contained a higher amount of Actinobacteriota and Firmicutes, phyla including several aerobic bacterial groups thriving in soil, as Streptomyces and Bacillus.

Concerning possible differences in nitrogen cycling by the rhizosphere bacterial community (which is partially related to anaerobic metabolism), variation in the abundance of members of Nitrospirota was found in Verna and Sieve varieties compared to that in Bologna and Andriolo varieties. Nitrospirota includes Nitrospira, a well-known genus of nitrite oxidizers present in agricultural soil and in the plant rhizosphere [45]. Interestingly, this result agrees with qPCR estimation of functional gene abundance. In fact, the analysis of genes involved in the nitrogen cycle showed that *amoA* (nitrification) and *nosZ* (denitrification) were differentially abundant with respect to varieties, suggesting that variety may lead to functional changes in the cycling of the plant-assimilable nitrogen. However, we could not a priori exclude fertilization may partially contribute to this difference. Experiments with bulk soil samples taken at higher distances from the plant may be helpful to clarify how much of the differences resulted from the sole fertilization. Indeed, it would highly relevant from the perspective of low-input precision agriculture to clarify the extent and impact on plant productivity of such variety–fertilization interaction in the rhizosphere bacterial community, to exploit the potentialities of variety-specific bacterial community. Multiple seasons and long-term field experiments in different soil and climatic conditions would be needed to fully address this question.

4. Materials and Methods

4.1. Experimental Field and Plant Genotypes

An experimental field was established in October 2019 at the "Grappi Luchino" organic farm (43.0507° N, 11.6912° E) close to Pienza (Tuscany, Italy). The soil was a moderately alkaline (pH 8.2) silty clay, with 10.5% sand, 49% silt, and 40.5% clay. The 0–40 cm layer (the most affected by either plant root apparatus and agronomic practices) contained 131 (3.1) g kg^{-1} total lime, 9.3 (0.2) g kg^{-1} total organic carbon, 1119 (10) mg kg^{-1} total nitrogen (N), 5.5 (0.3) mg kg^{-1} available phosphorus (P), and 189.2 (6) mg kg^{-1} available potassium (K). Cultivation of crops was performed under an organic farming system. Chickpea (*Cicer arietinum*, L.) was the previous crop. Residues from the previous crop were used as manure.

Four Italian varieties of winter wheat (*Triticum aestivum*, L.) were used in this study, including three "old" genotypes (namely Andriolo, Sieve, and Verna) [19,24,28], and one dwarf registered cultivar (namely Bologna). According to Migliorini et al., (2016), the germplasm of the four varieties has different origins: Verna is an awn-less cultivar established in 1953 from a breed between "Est Mottin 72" × "Mont Calme 245" [28]. Sieve is

an awn-less cultivar constituted in 1966 from "Est Mottin 72" × "Bellevue II" [46]. Andriolo is a Tuscan landrace having an unknown pedigree and dating back to the early 1900 [28]. Bologna is one of the most cultivated modern varieties in Italy owing to its quality and high yield stability; Bologna was bred in 1999 from ("H89092" × "H89136") × "Soissons" by Società Italiana Sementi (Italy) [47].

The experiment included 16 treatments (parcels), which were combinations of four varieties of winter wheat, two nitrogen (N) fertilization levels (40 and 80 kg N ha^{-1} yr^{-1}, N40 and N80, respectively), and two nitrogen-to-phosphoric anhydride (N:P_2O_5) fertilization ratios (NP) (N:P2O5 1:1 and 2:1, N1P1, and N2P1).

Soil tillage was carried out to a depth of 0.4 m with a moldboard plow in October 2019, followed by a tandem disk harrow (0.1 m depth) to break clods. The organic fertilizer "Endurance N8" (Unimer s.p.a., Milano, Italy), containing organic nitrogen from animal protein source, and the organic soft rock phosphate "GAFSA 27" (Panfertil s.p.a, Ravenna, Italy) were used as N and P_2O_5 sources, respectively. On 27 November, 2019, both fertilizers were broadcasted on the soil surface, immediately followed by spike tooth harrow (0.05 m depth) to incorporate the fertilized and prepare the false seedbed. Weeds were uprooted to a depth of 0.05 m using a spring tine harrow on 7 and 17 December 2019. Winter wheat seeds were sown 90 kg ha^{-1} on 17 December 2019 with a row spacing of 0.13 m. Grain harvesting was performed at crop commercial maturity (grain moisture lower than 13%) on 15 July 2020.

4.2. Sampling and Samples Treatment

Sampling of bulk soil (thereafter indicated as soil) and rhizosphere soil was performed in May 2020. For each plot, five individual plants at flowering stage were selected, separated at least 5 m from each other. After digging up the entire plant (10 cm radius around the plant, and down to 20 cm soil depth) and shaking off the plant roots, the soil remaining attached to the roots was detached with a sterile scalpel and used as rhizosphere soil. Bulk soil was collected using a stainless steel sterile corer to sample the top 10 cm of soil in patches of bare soil between plants (ca. 30 cm apart from the plant) and sieved at 2 mm. Bulk and rhizosphere soil samples for each plot were combined to create composite bulk and rhizosphere soil samples. A total of 32 combined samples from 16 plots were analyzed (Table 2). Samples were stored at −80 °C until DNA extraction.

Table 2. Samples analyzed. The code, wheat variety, the fertilization of the plot, and their geographical position is reported. Numbers in the sample code refer to the plots. "S" and "R" in the sample code indicate soil and rhizosphere fraction, respectively.

Variety	Code	Nitrogen Fertilization (N kg ha^{-1} yr^{-1})	Phosphorous Fertilization (P_2O_5 kg ha^{-1} yr^{-1})	Latitude (N)	Longitude (E)
Verna	V1S	40	80	43.051129°	11.690833°
	V1R	40	80	43.051129°	11.690833°
	V2S	80	160	43.050959°	11.691126°
	V2R	80	160	43.050959°	11.691126°
	V3S	80	80	43.050755°	11.691438°
	V3R	80	80	43.050755°	11.691438°
	V4S	40	40	43.050547°	11.691890°
	V4R	40	40	43.050547°	11.691890°
Sieve	S5S	40	40	43.050953°	11.692320°
	S5R	40	40	43.050953°	11.692320°
	S6S	80	80	43.051126°	11.691964°
	S6R	80	80	43.051126°	11.691964°
	S7S	80	160	43.051343°	11.691626°
	S7R	80	160	43.051343°	11.691626°
	S8S	40	80	43.051525°	11.691219°
	S8R	40	80	43.051525°	11.691219°

Table 2. *Cont.*

Variety	Code	Nitrogen Fertilization (N kg ha^{-1} yr^{-1})	Phosphorous Fertilization (P_2O_5 kg ha^{-1} yr^{-1})	Latitude (N)	Longitude (E)
Andriolo	A9S	40	40	43.049740°	11.690994°
	A9R	40	40	43.049740°	11.690994°
	A10S	80	80	43.049901°	11.690666°
	A10R	80	80	43.049901°	11.690666°
	A11S	80	160	43.050036°	11.690376°
	A11R	80	160	43.050036°	11.690376°
	A12S	40	80	43.050180°	11.690020°
	A12R	40	80	43.050180°	11.690020°
Bologna	B13S	40	80	43.050763°	11.690400°
	B13R	40	80	43.050763°	11.690400°
	B14S	80	160	43.050531°	11.690761°
	B14R	80	160	43.050531°	11.690761°
	B15S	80	80	43.050347°	11.691104°
	B15R	80	80	43.050347°	11.691104°
	B16S	40	40	43.050193°	11.691440°
	B16R	40	40	43.050193°	11.691440°

4.3. Soil Chemical and Physical Analyses

Soil pH was determined in a 1:5 soil-to-distilled water solution. Soil total nitrogen (STN) was determined by applying the dry combustion method on 5.0 mg soil samples, using a CHNS analyzer (CHN-S Flash E1112, Thermo-Finnigan LLC, San Jose, CA, USA) [48,49]. Soil samples (5.0 mg) were pretreated with chloride acid solution (10%) within silver cups to volatize the carbonate carbon and then analyzed by means CHNS to determine soil organic carbon (SOC) [50,51]. Ammonium nitrogen and nitrate nitrogen were determined by spectrophotometric determination (Lambda 20 spectrophotometer; PerkinElmer, Waltham, MA, USA) after extraction with the calcium chloride ($CaCl_2$) procedure by Houba et al., (1995), which has the advantage of extraction uniformity for the considered N forms [52]. Aliquots of soil extract solution were treated according to the nitrate copper-cadmium reduction method [53], and the resulting solutions were spectrophotometrically analyzed at 540 nm after Griess reaction [54] for the determination of nitrate nitrogen. Further, aliquots of soil extract solution were treated following the Nessler method [55] and then absorbance analyzed at 420 nm for the determination of ammonium nitrogen. According to [56], soil samples were extracted with 0.5 M $NaHCO_3$ at pH 8.5 to solubilize inorganic P forms which were spectrophotometrically analyzed at 882 nm after reaction with ammonium molybdate to determine the soil assimilable phosphorous (AP). Exchangeable cations (calcium, iron, magnesium, potassium, selenium, sodium, and zinc) were extracted from the soil using a 1 N NH_4OAc extracting solution pH 7.0 and then determined using an ICP-OES analyzer (ICAP™ 7400 ICP-OES Analyzer, Thermo Scientific, Waltham, MA, USA). The soil cation exchange capacity (CEC) was determined using the BaCl2-triethanolamine method as proposed by Mehlich (1939) [57].

4.4. eDNA Extraction and 16S rRNA Gene Amplicon Sequencing

Environmental DNA (eDNA) was extracted from 500 mg of soil using a DNeasy PowerSoil Pro (Qiagen). From the extracted DNA, the bacterial V4 region of 16S rRNA genes was amplified with primers 515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806R (5′-GACTACHVGGGTATCTAATCC-3′) [58] in a 25 µL total volume with KAPA HiFi HotStart ReadyMix, 1 µM of each primer, and 10 ng of template DNA with 25 cycles with the following temperature profile: 30 s at 95 °C, 30 s at 55 °C, and 30 s at 72 °C. PCR products were sequenced in a single run using Illumina MiSeq technology with pair-end sequencing strategy and a MiSeq Reagent Kit v3 (Illumina, USA). Library preparation

(Nextera XT, Illumina, San Diego, CA, USA) and demultiplexing were performed following Illumina's standard pipeline as previously reported [10].

4.5. Quantitative PCR of Genes Involved in the Nitrogen Cycle

Quantitative PCR was performed on eDNA using primers for the amplification of *nifH*, *amoA*, *nosZ*, *nirK* genes (primer sequences and citations are reported in Table S8). Total bacterial load was estimated on the 16S rRNA gene. Reactions were performed with 5 ng of template DNA in 10 µL total volume on a QuantStudio7 Flexi apparatus (Applied Biosystems) with Sybr Green technology (Maxima SYBR Green qPCR Master Mix, Thermofisher). Reactions were performed in triplicates. Relative abundance of *nifH*, *amoA*, *nosZ*, *nirK* genes was expressed in relation to 16SrRNA, as previously reported [59,60].

4.6. Bioinformatic and Statistical Analyses

Illumina reads were trimmed using the "Trim Galore!" tool on Galaxy server (https://usegalaxy.org/, last accessed on 21 December 2021). Paired-end sequences were clustered into Amplicon Sequence Variants (ASVs) following the DADA2 pipeline (version 1.16) [61]. After filtering the sequences and removing the chimeras, the taxonomy assignment was carried out comparing our data against the SILVA NR99rel138 standard database of bacteria [62] using "DECIPHER" R package (version 2.18.1) [63] as implementation of DADA2 (SSU version 138 available at: http://www2.decipher.codes/Downloads.html, last accessed on 21 December 2021). Annotated ASVs count tables were processed in Phyloseq package in R environment version 4.0.5 [64].

The analysis of microbial communities was performed through the "Phyloseq" R package (version 1.34.0) [64]. For alpha diversity analysis, the Shannon and Simpson indices were calculated and plotted using the function "diversity()" within "microbiome" R package (version 1.12.0) [65]. Good's coverage and Evenness indices were calculated through the R functions "goods()" and "evenness()", respectively, within the "microbiome" R package (version 1.12.0). A Wilcoxon test for multiple comparison of averages was performed on alpha diversity indices using the "ggviolin()" and "stat_compare_means()" within the "ggpubr" R package (version 0.4.0). For ordination plots of phyloseq objects, a multivariate analysis based on Bray–Curtis distance and NMDS ordination generated using the "ordinate" function was performed, and plots were generated using the "plot_ordination()" function within phyloseq package. Rarefaction curves were generated using the "ggplot2" (version 3.3.3) [66] and "ranacapa" (version 0.1.0) [67] R packages using the "ggrare()" function on the phyloseq object. The "ggplot2" R package (version 3.3.3) was used to generate relative abundance plots.

Different community structures were analyzed using permutational multivariate analysis of variance (PERMANOVA) performed using the R packages "ggplot2" (version 3.3.3), "vegan" (version 2.5-7), and "pairwise.Adonis" (version 0.0.1) using the functions "adonis2()" and "pairwise.adonis()", respectively. Clustering analysis for relationship discovery was performed using the "hclust()" R function selecting the agglomeration method "complete" on the distance matrix produced by the "dist()" function within the "stats" R package (version 4.0.5).

Statistical analyses for quantitative PCR data were performed in R environment (version 4.0.5) calculating the ratio between the threshold cycles of the target genes and 16S gene. Graphics of the ratio averages were performed using the "ggboxplot()" function within the "ggplot2′ R package (version 3.3.3). A Scott–Knott test was performed on averages using the functions "with()" and "SK()" within the "ScottKnott" R package (version 1.2-7) [68].

4.7. Prediction of Functional Abundances

The software PICRUSt2 (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States, https://github.com/picrust/picrust2, last accessed on 21 December 2021) was utilized to investigate and predict the functional potential for each determined

sequence in the microbial community. PICRUSt2 pipeline returned as output the metabolic pathway abundances, and functional gene profiling by comparison with MetaCyc and Kyoto Encyclopedia of Genes and Genomes databases [69]. Ranking of the most relevant pathways for varieties differentiation was conducted by running a SIMPER test using the function "simper()" with the "vegan" R package (version 2.5-7). For each contrast between varieties, the top three differentially expressed pathways were selected, and analyses were carried out using the function "metaMDS()" in the "vegan" R package (version 2.5-7) to perform nonmetric multidimensional scaling (nMDS) and the function "heatmap()" in the "stats" R package (version 4.0.5) to graphically represent results.

4.8. Links to Deposited Data

The sequences dataset was deposited in the SRA database under the BioProject PRJNA720503 (last accessed on 21 December 2021).

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms23073616/s1.

Author Contributions: Conceptualization, M.M., M.N., S.O., A.M. and L.C.; methodology, L.C.; M.N.; validation, M.N., M.M., L.C., C.F. and F.V.; formal analysis, L.C.; investigation, L.C., C.F., F.V. and M.N.; resources, M.N., M.M., S.O. and A.M.; data curation, A.M., L.C.; writing—original draft preparation, A.M., L.C. and M.N.; writing—review and editing, M.N., M.M., S.O., A.M., L.C. and C.F.; supervision, A.M. and L.C.; project administration, A.M.; funding acquisition, A.M., S.O. and M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are available at https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA720503, accessed on 1 February 2022.

Acknowledgments: This work was partially supported by the grant MICRO4Legumes, D.M.n.89267 (Italian Ministry of Agriculture) to A.M. The project was part of Measure 16.2 "FACS Frumenti Antichi e Coltivazioni Sostenibili", founded by the 2014–2020 Rural Development Programme for Tuscany. The authors want to thank Roberto Vivoli, Carolina Fabbri, and Stefania Truschi of DAGRI for their technical support. A special thank is given to "Azienda Agricola Grappi Luchino".

Conflicts of Interest: The authors declare no conflict of interests.

References

1. Bulgarelli, D.; Schlaeppi, K.; Spaepen, S.; Van Themaat, E.V.L.; Schulze-Lefert, P. Structure and functions of the bacterial microbiota of plants. *Annu. Rev. Plant Biol.* **2013**, *64*, 807–838. [CrossRef] [PubMed]
2. Schlaeppi, K.; Bulgarelli, D. The Plant Microbiome at Work. *Mol. Plant-Microbe Interact. MPMI* **2015**, *212*, 212–217. [CrossRef] [PubMed]
3. Compant, S.; Samad, A.; Faist, H.; Sessitsch, A. A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. *J. Adv. Res.* **2019**, *19*, 29–37. [CrossRef] [PubMed]
4. Sánchez-Cañizares, C.; Jorrín, B.; Poole, P.S.; Tkacz, A. Understanding the holobiont: The interdependence of plants and their microbiome. *Curr. Opin. Microbiol.* **2017**, *38*, 188–196. [CrossRef]
5. Vandenkoornhuyse, P.; Quaiser, A.; Duhamel, M.; Le Van, A.; Dufresne, A. The importance of the microbiome of the plant holobiont. *New Phytol.* **2015**, *206*, 1196–1206. [CrossRef]
6. Agler, M.T.; Ruhe, J.; Kroll, S.; Morhenn, C.; Kim, S.T.; Weigel, D.; Kemen, E.M. Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. *PLoS Biol.* **2016**, *14*, e1002352. [CrossRef]
7. Martínez-Romero, E.; Aguirre-Noyola, J.L.; Taco-Taype, N.; Martínez-Romero, J.; Zuñiga-Dávila, D. Plant microbiota modified by plant domestication. *Syst. Appl. Microbiol.* **2020**, *43*, 126106. [CrossRef]
8. Bulgarelli, D.; Garrido-Oter, R.; Münch, P.C.; Weiman, A.; Dröge, J.; Pan, Y.; McHardy, A.C.; Schulze-Lefert, P. Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* **2015**, *17*, 392–403. [CrossRef]
9. Escudero-Martinez, C.; Bulgarelli, D. Tracing the evolutionary routes of plant–microbiota interactions. *Curr. Opin. Microbiol.* **2019**, *49*, 34–40. [CrossRef]

10. Checcucci, A.; Maida, I.; Bacci, G.; Ninno, C.; Bilia, A.R.; Biffi, S.; Firenzuoli, F.; Flamini, G.; Fani, R.; Mengoni, A. Is the plant-associated microbiota of Thymus spp. adapted to plant essential oil? *Res. Microbiol.* **2017**, *168*, 276–282. [CrossRef]
11. Kristin, A.; Miranda, H. The root microbiota—A fingerprint in the soil? *Plant Soil* **2013**, 671–686. [CrossRef]
12. Fagorzi, C.; Bacci, G.; Huang, R.; Cangioli, L.; Checcucci, A.; Fini, M.; Perrin, E.; Natali, C.; diCenzo, G.C.; Mengoni, A. Nonadditive transcriptomic signatures of genotype-by-genotype interactions during the initiation of plant-rhizobium symbiosis. *mSystems* **2021**, *6*, 00974-20. [CrossRef] [PubMed]
13. Busby, P.E.; Soman, C.; Wagner, M.R.; Friesen, M.L.; Kremer, J.; Bennett, A.; Morsy, M.; Eisen, J.A.; Leach, J.E.; Dangl, J.L. Research priorities for harnessing plant microbiomes in sustainable agriculture. *PLoS Biol.* **2017**, *15*, e2001793. [CrossRef] [PubMed]
14. Stopnisek, N.; Shade, A. Prioritizing persistent microbiome members in the common bean rhizosphere: An integrated analysis of space, time, and plant genotype. *ISME J.* **2021**, *15*, 2708–2722. [CrossRef] [PubMed]
15. Poole, N.; Donovan, J.; Erenstein, O. Viewpoint: Agri-nutrition research: Revisiting the contribution of maize and wheat to human nutrition and health. *Food Policy* **2021**, *100*, 101976. [CrossRef]
16. Fróna, D.; Szenderák, J.; Harangi-Rákos, M. The Challenge of Feeding the World. *Sustainability* **2019**, *11*, 5816. [CrossRef]
17. OECD-FAO. *OECD-FAO Agricultural Outlook 2017-2026*; OECD-FAO Agricultural Outlook; OECD: Paris, France, 2017.
18. Herrera, J.M.; Levy Häner, L.; Mascher, F.; Hiltbrunner, J.; Fossati, D.; Brabant, C.; Charles, R.; Pellet, D. Lessons from 20 Years of Studies of Wheat Genotypes in Multiple Environments and under Contrasting Production Systems. *Front. Plant Sci.* **2020**, *10*, 1745. [CrossRef]
19. Guerrini, L.; Napoli, M.; Mancini, M.; Masella, P.; Cappelli, A.; Parenti, A.; Orlandini, S. Wheat grain composition, dough rheology and bread quality as affected by nitrogen and sulfur fertilization and seeding density. *Agronomy* **2020**, *10*, 233. [CrossRef]
20. Kavamura, V.N.; Robinson, R.J.; Hughes, D.; Clark, I.; Rossmann, M.; Soares de Melo, I.; Hirsch, P.R.; Mendes, R.; Mauchline, T.H. Wheat Dwarfing Influences Selection of the Rhizosphere Microbiome. *Sci. Rep.* **2020**, *10*, 1452. [CrossRef]
21. Ellis, M.; Rebetzke, G.; Chandler, P.; Bonnett, D.; Spielmeyer, W.; Richards, R. The effect of different height reducing genes on the early growth of wheat. *Funct. Plant Biol.* **2004**, *31*, 583–589. [CrossRef]
22. Arshad, M.A. A Review on Wheat Management, Strategies, Current Problems and Future Perspectives. *Haya Saudi J. Life Sci.* **2021**, *6*, 14–18. [CrossRef]
23. Sanchez-Garcia, M.; Álvaro, F.; Peremarti, A.; Martín-Sánchez, J.A.; Royo, C. Changes in bread-making quality attributes of bread wheat varieties cultivated in Spain during the 20th century. *Eur. J. Agron.* **2015**, *63*, 79–88. [CrossRef]
24. Migliorini, P.; Spagnolo, S.; Torri, L.; Arnoulet, M.; Lazzerini, G.; Ceccarelli, S. Agronomic and quality characteristics of old, modern and mixture wheat varieties and landraces for organic bread chain in diverse environments of northern Italy. *Eur. J. Agron.* **2016**, *79*, 131–141. [CrossRef]
25. Fan, M.S.; Zhao, F.J.; Fairweather-Tait, S.J.; Poulton, P.R.; Dunham, S.J.; McGrath, S.P. Evidence of decreasing mineral density in wheat grain over the last 160 years. *J. Trace Elem. Med. Biol.* **2008**, *22*, 315–324. [CrossRef] [PubMed]
26. Shewry, P.R.; Hey, S. Do "ancient" wheat species differ from modern bread wheat in their contents of bioactive components? *J. Cereal Sci.* **2015**, *65*, 236–243. [CrossRef]
27. Arzani, A.; Ashraf, M. Cultivated Ancient Wheats (*Triticum* spp.): A Potential Source of Health-Beneficial Food Products. *Compr. Rev. Food Sci. Food Saf.* **2017**, *16*, 477–488. [CrossRef]
28. Ghiselli, L.; Rossi, E.; Whittaker, A.; Dinelli, G.; Baglio, A.P.; Andrenelli, L.; Benedettelli, S. Nutritional characteristics of ancient Tuscan varieties of *Triticum aestivum* L. *Ital. J. Agron.* **2016**, *11*, 237–245. [CrossRef]
29. Donn, S.; Kirkegaard, J.A.; Perera, G.; Richardson, A.E.; Watt, M. Evolution of bacterial communities in the wheat crop rhizosphere. *Environ. Microbiol.* **2015**, *17*, 610–621. [CrossRef]
30. Wolinska, A.; Kuzniar, A.; Galazka, A. Biodiversity in the Rhizosphere of selected winter wheat (*Triticum aestivum* L.) cultivars-genetic and catabolic fingerprinting. *Agronomy* **2020**, *10*, 953. [CrossRef]
31. Simonin, M.; Dasilva, C.; Terzi, V.; Ngonkeu, E.L.M.; DIouf, D.; Kane, A.; Béna, G.; Moulin, L. Influence of plant genotype and soil on the wheat rhizosphere microbiome: Evidences for a core microbiome across eight African and European soils. *FEMS Microbiol. Ecol.* **2020**, *96*, fiaa067. [CrossRef]
32. Chen, S.; Waghmode, T.R.; Sun, R.; Kuramae, E.E.; Hu, C.; Liu, B. Root-associated microbiomes of wheat under the combined effect of plant development and nitrogen fertilization. *Microbiome* **2019**, *7*, 136. [CrossRef] [PubMed]
33. Tkacz, A.; Pini, F.; Turner, T.R.; Bestion, E.; Simmonds, J.; Howell, P.; Greenland, A.; Cheema, J.; Emms, D.M.; Uauy, C.; et al. Agricultural Selection of Wheat Has Been Shaped by Plant-Microbe Interactions. *Front. Microbiol.* **2020**, *11*, 132. [CrossRef] [PubMed]
34. Spor, A.; Roucou, A.; Mounier, A.; Bru, D.; Breuil, M.C.; Fort, F.; Vile, D.; Roumet, P.; Philippot, L.; Violle, C. Domestication-driven changes in plant traits associated with changes in the assembly of the rhizosphere microbiota in tetraploid wheat. *Sci. Rep.* **2020**, *10*, 12234. [CrossRef]
35. Germida, J.; Siciliano, S. Taxonomic diversity of bacteria associated with the roots of modern, recent and ancient wheat cultivars. *Biol. Fertil. Soils* **2001**, *33*, 410–415. [CrossRef]
36. Evenson, R.E.; Gollin, D. Assessing the impact of the Green Revolution, 1960 to 2000. *Science* **2003**, *300*, 758–762. [CrossRef]
37. Fan, K.; Delgado-Baquerizo, M.; Guo, X.; Wang, D.; Wu, Y.; Zhu, M.; Yu, W.; Yao, H.; Zhu, Y.G.; Chu, H. Suppressed N fixation and diazotrophs after four decades of fertilization. *Microbiome* **2019**, *7*, 1–10. [CrossRef]

38. Iannucci, A.; Canfora, L.; Nigro, F.; De Vita, P.; Beleggia, R. Relationships between root morphology, root exudate compounds and rhizosphere microbial community in durum wheat. *Appl. Soil Ecol.* **2021**, *158*, 103781. [CrossRef]
39. Mamine, F.; Farès, M.H. Barriers and levers to developing wheat–pea intercropping in Europe: A review. *Sustainability* **2020**, *12*, 6962. [CrossRef]
40. Lee, S.H.; Ka, J.O.; Cho, J.C. Members of the phylum Acidobacteria are dominant and metabolically active in rhizosphere soil. *FEMS Microbiol. Lett.* **2008**, *285*, 263–269. [CrossRef]
41. Deng, S.; Caddell, D.F.; Xu, G.; Dahlen, L.; Washington, L.; Yang, J.; Coleman-Derr, D. Genome wide association study reveals plant loci controlling heritability of the rhizosphere microbiome. *ISME J.* **2021**, *15*, 3181–3194. [CrossRef]
42. Muhammad, I.; Yang, L.; Ahmad, S.; Zeeshan, M.; Farooq, S.; Ali, I.; Khan, A.; Zhou, X.B. Irrigation and nitrogen fertilization alter soil bacterial communities, soil enzyme activities, and nutrient availability in maize crop. *Front. Microbiol.* **2022**, *105*, 833758. [CrossRef] [PubMed]
43. Ossowicki, A.; Tracanna, V.; Petrus, M.L.C.; van Wezel, G.; Raaijmakers, J.M.; Medema, M.H.; Garbeva, P. Microbial and volatile profiling of soils suppressive to Fusarium culmorum of wheat. *Proc. R. Soc. B Biol. Sci.* **2020**, *287*, 20192527. [CrossRef] [PubMed]
44. Cai, Y.; Shen, J.P.; Di, H.J.; Zhang, L.M.; Zhang, C.; He, J.Z. Variation of soil nitrate and bacterial diversity along soil profiles in manure disposal maize field and adjacent woodland. *J. Soils Sediments* **2020**, *20*, 3557–3568. [CrossRef]
45. Wang, X.; Wang, S.; Jiang, Y.; Zhou, J.; Han, C.; Zhu, G. Comammox bacterial abundance, activity, and contribution in agricultural rhizosphere soils. *Sci. Total Environ.* **2020**, *727*, 138563. [CrossRef] [PubMed]
46. Wilhelmi, K.D.; Kuhr, S.L.; Johnson, V.A.; Mattern, P.J.; Schmidt, J.W. *Results of the Seventh International Winter Wheat Performance Nursery Grown in 1975*; ral Resources, Agricultural Experiment Station: Lincoln, NB, USA, 1977.
47. Verrillo, F.; Badeck, F.-W.; Terzi, V.; Rizza, F.; Bernardo, L.; Di Maro, A.; Fares, C.; Zaldei, A.; Miglietta, F.; Moschella, A.; et al. Elevated field atmospheric CO2 concentrations affect the characteristics of winter wheat (cv. Bologna) grains. *Crop Pasture Sci.* **2017**, *68*, 713–725. [CrossRef]
48. Jimenez, R.R.; Ladha, J.K. Automated elemental analysis: A rapid and reliable but expensive measurement of total carbon and nitrogen in plant and soil samples. *Commun. Soil Sci. Plant Anal.* **2008**, *24*, 1897–1924. [CrossRef]
49. Matejovic, I. Determination of carbon and nitrogen in samples of various soils by the dry combustion. *Commun. Soil Sci. Plant Anal.* **1997**, *28*, 1499–1511. [CrossRef]
50. Vitti, C.; Stellacci, A.M.; Leogrande, R.; Mastrangelo, M.; Cazzato, E.; Ventrella, D. Assessment of organic carbon in soils: A comparison between the Springer–Klee wet digestion and the dry combustion methods in Mediterranean soils (Southern Italy). *CATENA* **2016**, *137*, 113–119. [CrossRef]
51. Verardo, D.J.; Froelich, P.N.; McIntyre, A. Determination of organic carbon and nitrogen in marine sediments using the Carlo Erba NA-1500 analyzer. *Deep Sea Res. Part A. Oceanogr. Res. Pap.* **1990**, *37*, 157–165. [CrossRef]
52. Houba, V.J.G.; Huijbregts, A.W.M.; Wilting, P.; Novozamsky, I.; Gort, G. Sugar yield, nitrogen uptake by sugar beet and optimal nitrogen fertilization in relation to nitrogen soil analyses and several additional factors. *Biol. Fertil. Soils* **1995**, *19*, 55–59. [CrossRef]
53. *APHA*; APHA Standard Methods, Method 4500-NO3-E—Cadmium Reduction Method. 22nd ed. America Public Health Association: Washington, DC, USA, 2000.
54. *EPA EPA Method 353.2*; Determination of Nitrate-Nitrite Nitrogen by Automated Colorimetry. Revision 2.0. August 1993. United States Environmental Protection Agency: Cincinnati, OH, USA, 19 August 1993.
55. *ASTM D1426-15*; Standard Test Methods for Ammonia Nitrogen in Water. ASTM International: West Conshohocken, PA, USA, 2015.
56. Olsen, S.R.; Sommers, L.E. Phosphorus. In *Methods of Soil Analysis Part 2 Chemical and Microbiological Properties*; Page, A.L., Ed.; American Society of Agronomy, Soil Science Society of America: Madison, WI, USA, 1982; pp. 403–430.
57. Mehlich, A. Use of Triethanolamine Acetate-Barium Hydroxide Buffer for the Determination of Some Base Exchange Properties and Lime Requirement of Soil. *Soil Sci. Soc. Am. J.* **1939**, *3*, 162–166. [CrossRef]
58. Klindworth, A.; Pruesse, E.; Schweer, T.; Peplies, J.; Quast, C.; Horn, M.; Glöckner, F.O. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **2013**, *41*, e1. [CrossRef] [PubMed]
59. Russo, E.; Bacci, G.; Chiellini, C.; Fagorzi, C.; Niccolai, E.; Taddei, A.; Ricci, F.; Ringressi, M.N.; Borrelli, R.; Melli, F.; et al. Preliminary comparison of oral and intestinal human microbiota in patients with colorectal cancer: A pilot study. *Front. Microbiol.* **2018**, *8*, 2699. [CrossRef] [PubMed]
60. Abdelrhman, K.F.A.; Bacci, G.; Nistri, A.; Mengoni, A.; Ugolini, A. Diet and gut microbiota of two supralittoral amphipods Orchestia montagui and Talitrus saltator living in different microhabitats. *Estuar. Coast. Shelf Sci.* **2017**, *197*, 119–125. [CrossRef]
61. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13*, 581–583. [CrossRef]
62. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, D590–D596. [CrossRef]
63. Wright, E.S.; Yilmaz, L.S.; Noguera, D.R. DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. *Appl. Environ. Microbiol.* **2012**, *78*, 717–725. [CrossRef]
64. McMurdie, P.J.; Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **2013**, *8*, e61217. [CrossRef]
65. Shetty, S.A.; Lahti, L. Microbiome data science. *J. Biosci.* **2019**, *44*, 115. [CrossRef]

66. *Wickham Hadley ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2009; ISBN 978-0-387-98140-6.
67. Kandlikar, G.S.; Gold, Z.J.; Cowen, M.C.; Meyer, R.S.; Freise, A.C.; Kraft, N.J.B.; Moberg-Parker, J.; Sprague, J.; Kushner, D.J.; Curd, E.E. ranacapa: An R package and Shiny web app to explore environmental DNA data with exploratory statistics and interactive visualizations. *F1000Research* **2018**, *7*, 1734. [CrossRef]
68. Jelihovschi, E.; Faria, J.C.; Allaman, I.B.; Allaman, M.I.B. ScottKnott: A package for performing the Scott-Knott clustering algorithm in R. *TEMA* **2014**, *15*, 3–17. [CrossRef]
69. Douglas, G.M.; Maffei, V.J.; Zaneveld, J.; Yurgel, S.N.; Brown, J.R.; Taylor, C.M.; Huttenhower, C.; Langille, M.G.I. PICRUSt2: For prediction of metagenome functions. *Nat Biotechnol.* **2020**, *38*, 685–688. [CrossRef] [PubMed]

microorganisms

MDPI

Article

Effect of Site and Phenological Status on the Potato Bacterial Rhizomicrobiota

Lisa Cangioli [1], Marco Mancini [2], Ada Baldi [2,*], Camilla Fagorzi [1], Simone Orlandini [2], Francesca Vaccaro [1] and Alessio Mengoni [1,*]

[1] Department of Biology, University of Florence, Via Madonna del Piano 6, 50019 Sesto Fiorentino, Italy
[2] Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, Piazzale delle Cascine 18, 50144 Firenze, Italy
* Correspondence: ada.baldi@unifi.it (A.B.); alessio.mengoni@unifi.it (A.M.)

Abstract: The potato is the fourth major food crop in the world. Its cultivation can encounter problems, resulting in poor growth and reduced yield. Plant microbiota has shown an ability to increase growth and resistance. However, in the development of effective microbiota manipulation strategies, it is essential to know the effect of environmental variables on microbiota composition and function. Here, we aimed to identify the differential impact of the site of cultivation and plant growth stage on potato rhizosphere microbiota. We performed a 16S rRNA gene amplicon sequencing analysis of rhizospheric soil collected from potato plants grown at four sites in central Italy during two phenological stages. Rhizomicrobiota was mainly composed of members of phyla Acidobacteriota, Actinobacteriota, Chloroflexi, and Proteobacteria and was affected by both the site of cultivation and the plant stages. However, cultivation sites overcome the effect of plant phenological stages. The PiCRUST analysis suggested a high abundance of functions related to the biosynthesis of the siderophore enterobactin. The presence of site-specific taxa and functional profiling of the microbiota could be further exploited in long-term studies to evaluate the possibility of developing biomarkers for traceability of the products and to exploit plant growth-promoting abilities in the native potato microbiota.

Keywords: potato; rhizosphere; microbiota; ecosystem functions

check for updates

Citation: Cangioli, L.; Mancini, M.; Baldi, A.; Fagorzi, C.; Orlandini, S.; Vaccaro, F.; Mengoni, A. Effect of Site and Phenological Status on the Potato Bacterial Rhizomicrobiota. *Microorganisms* **2022**, *10*, 1743. https://doi.org/10.3390/microorganisms10091743

Academic Editor: Attile Kereszt

Received: 16 August 2022
Accepted: 26 August 2022
Published: 29 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The potato is the fourth major food crop in the world, and its cultivation in the field can encounter many problems that result in poor growth and reduced yield [1]. Aside from plant genetics, crop yield is the result of a variety of factors, including agronomic practices, soil, climatic conditions, and plant microbiota [2]. The plant microbiota, and especially the rhizospheric microbiota, has been clearly related to crop yield and resistance to biotic and abiotic stresses [3–7]. As for many other plant species (see, for instance, [8]), the microbiota of potato rhizosphere has been shown to be strongly influenced by the plant developmental stage [9–12], with the flowering stage generally having the highest diversity [9,10]. However, the effect of the site of cultivation is also relevant, since pedoclimatic conditions affect both soil microbiota composition and plant growth. In a survey of rhizosphere soil samples from potatoes grown at different altitudes [9], it was shown that richness (alpha diversity) was more affected by the site of sampling, as conditioned by different soil properties and climatic conditions, while the taxonomic composition was more significantly affected by the plant vegetation stage. When considering the exploitation of the plant microbiota for improving plant growth, in terms of the preparation and formulation of bioinoculants, it is vital to identify the core set of microbial taxa associated with the target plant species—i.e., those which are always present, irrespective of environmental changes (e.g., the site of cultivation, growth stage, agronomic practices) and, possibly, of plant variety [13]. At the same time, the identification of taxa that can vary depending on the above-mentioned

variables can permit the design of tailored bioinoculant formulations that are optimized under specific conditions [14,15]. It can also define possible microbiological markers of, for instance, the locality of cultivation, to guarantee traceability and enhance the area of the origin of the product [16–18]. This latter aspect is particularly relevant for Italy. In fact, although Italy is not one of the main potato producers [19], there are several traditional dishes that use local varieties of potatoes—especially sites suitable for potato cultivation [20–22] in the hilly and mountainous areas of the Central Apennines.

Consequently, it is highly relevant to perform an investigation of potato-associated microbiota to provide additional data for a better understanding of the role phenological stages and cultivation sites play on potato microbiota diversity and, consequently, future rational development of bioinoculants and microbiome fingerprinting methods.

In this work, we aimed to identify the components of the bacterial rhizospheric microbiota of potatoes and their changes in relation to the site of cultivation and the plant growth stage on cv. Kennebec, a classical potato variety grown in the Italian Central Apennine [23]. We answered this question by performing a 16S rRNA gene amplicon sequencing analysis of rhizospheric soil collected from sites differing in pedoclimatic conditions during two phenological stages: flowering and harvesting.

2. Materials and Methods

2.1. Potato Field Selection, Sampling, and Physicochemical Analysis

Four sampling sites cultivated with potatoes (*Solanum tuberosum* L.), cv. Kennebec, were selected in Tuscany (central Italy) at different altitudes (Table 1). For the locality of Gavigno, two sampling sites were considered for the evaluation of differences over a small distance (ca. 650 m distance). Fields were tilled and fertilized (Table 1) before sowing. and sites did not receive additional irrigation. For each sampling site, three soil samples were collected at the time of sowing, at a depth of 15 cm, and then mixed and analyzed to assess the main physical and chemical properties. The climatic conditions were computed for the entire growing cycle, from approximately early April 2020 (sowing) to late August 2020 (maturity). The temperature pattern was computed using growing degree days (GDDs) following the method suggested by [24]:

$$\sum \text{GDD} = \frac{(\text{Tmax} + \text{Tmin})}{2} - \text{Tbase} \quad (1)$$

where ∑GDD is the sum of the degree days, Tmax is the maximum air temperature (°C), Tmin is the minimum air temperature (°C), and Tbase is the lower baseline temperature of the crop (°C). In this work, a Tbase of 4.4 °C was used [25]. The amount of water supplied by rainfalls was calculated as cumulated monthly precipitation.

Table 1. Sampling sites analyzed in this study. The locality; geographical coordinates; and nitrogen (N), phosphorus (P), and potassium (K) fertilization rates are reported.

Locality	Geographical Coordinates	Altitude	NPK Fertilization Rate
Gavigno 1	44.068862° N, 11.093375° E	739 m	15-5-5 (500 g ha^{-1})
Gavigno 2	44.062105° N, 11.095891° E	763 m	15-5-5 (500 kg ha^{-1})
Val di Chiana	43.412705° N, 11.784870° E	260 m	15-9-15 (100 kg ha^{-1})
Prato	43.860192° N, 11.054822° E	39 m	12-12-17 (350 kg ha^{-1})

The sampling of soil was performed during the flowering and harvesting stages, from June to August 2020, under the guidance of the sampling campaign for the Crop Microbiome Survey (https://www.globalsustainableagriculture.org/the-crop-microbiome-survey/) [26], following the protocol defined in March 2020 (https://www.globalsustainableagriculture.org/wp-content/uploads/CropMicrobiome_March-2020.pdf, accessed on 1 June 2020). For each sampling site and time, six samples were prepared, for a total of 48 samples (Table S1). Each sample was constituted of five individual plants at the flower-

ing stage or the harvesting stage. Selected plants were at a distance of at least 5 m from each other. After digging up the entire plant (10 cm radius around the plant and down to 20 cm soil depth) and shaking off the plant roots, the soil remaining attached to the roots was detached with a sterile scalpel and considered to be rhizospheric soil. Rhizospheric soils from each of the five plants were mixed, vortexed, and stored at −80 °C until DNA extraction. Physicochemical analyses of the soils were performed as reported previously [27].

2.2. eDNA Extraction and 16S rRNA Gene Amplicon Sequencing

Environmental DNA (eDNA) was extracted from 500 mg of soil using the DNeasy PowerSoil Pro (Qiagen Italy, Milan, Italy). From the extracted DNA, the bacterial V4 region of the 16S rRNA genes was amplified, as reported previously, from 1 ng of total DNA [27]. Primers 515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806R (5′-GACTACHVGGGTATCTAATCC-3′) were used [28]. Reactions were performed in a 25 µL total volume with KAPA HiFi HotStart ReadyMix, (Kapabiosystems, Cape Town, South Africa) 1 µM of each primer with 25 cycles, with the following temperature profile: 30″ 95 °C, 30″ 55 °C, and 30″ 72 °C. PCR products were sequenced in a single run using Illumina MiSeq technology with a pair-end sequencing strategy and a MiSeq Reagent Kit v3 (Illumina, San Diego, CA, USA). Library preparation (Nextera XT, Illumina, San Diego, CA, USA) and demultiplexing were performed following Illumina's standard pipeline, as previously reported [29].

2.3. Raw Data Processing, Clustering, and Taxonomic Assignment of Reads

Illumina reads were trimmed using the 'Trim Galore!' tool on the Galaxy server (https://usegalaxy.org/, accessed on 10 October 2021). Paired-end sequences were clustered into amplicon sequence variants (ASVs), following the DADA2 pipeline (version 1.16) [30]. Sequences were filtered and dereplicated using default parameters (MaxN = 0, truncQ = 2, rm.phix = TRUE, and maxEE = 2) to collapse identical reads into unique sequences, and chimeras were removed. The taxonomy assignment was carried out by comparing our data against the SILVA NR99rel138 standard database of bacteria [31], using the 'DECIPHER' R package (version 2.18.1) [32] as an implementation of DADA2 (SSU version 138 available at: http://www2.decipher.codes/Downloads.html, accessed on 20 October 2021). Annotated ASV count tables were processed using the Phyloseq package in R environment version 4.0.5 [33].

2.4. Diversity of the Rhizosphere Microbiota

The analysis of the microbial communities was mainly performed by following a previously published pipeline [27] through the 'Phyloseq' R package (version 1.34.0) [33]. For alpha diversity analyses, the Shannon and Simpson indices were calculated and plotted using the function 'diversity()' within the 'microbiome' R package (version 1.12.0) [34]. Good's coverage and evenness indices were calculated through the R functions 'goods()' and 'evenness()', respectively, within the 'microbiome' R package (version 1.12.0). Two evenness indices were computed (Camargo's index [35] and Pielou's index [36]). Rarefaction curves were generated using the 'ggplot2′ (version 3.3.3) [37] and 'ranacapa' (version 0.1.0) [38] R packages by using the 'ggrare' function on the phyloseq object. The 'ggplot2′ R package (version 3.3.3) was used to generate relative abundance plots. Wilcoxon tests for multiple comparisons of averages were performed on alpha diversity indices using 'ggviolin()' and 'stat_compare_means()' within the 'ggpubr' R package (version 0.4.0).

2.5. Effect of Site and Phenological Stage on Rhizosphere Taxonomic Composition

For the ordination plots of the phyloseq objects, a multivariate analysis based on the Bray–Curtis distance and nonmetric multidimensional scaling (nMDS) ordination (generated using the "ordinate" function) was performed, and plots were generated using the "plot_ordination()" function within the phyloseq package. Different community structures were analyzed using permutational a multivariate analysis of variance (PERMANOVA), which was performed using the R packages 'ggplot2′ (version 3.3.3), 'vegan' (version 2.5-7),

and 'pairwise.Adonis' (version 0.0.1), using the functions 'adonis2()' and 'pairwise.adonis()', respectively.

A differential abundance analysis was carried out using the R package DeSeq2 v1.30.1 [39], in order to identify the ASVs/taxa differentially expressed along the samples. The functions 'phyloseq_to_deseq2()', 'DESeq()', and 'results()' were used.

2.6. Functional Inference on the Rhizosphere Microbiota

The metabolic pathway abundances and functional gene profiles were obtained by running the software PICRUSt2 (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States, https://github.com/picrust/picrust2, accessed on 20 October 2021), as previously reported in [27], using the Galaxy pipeline of the Deng Lab [40]. The ranking of the most relevant pathways for variety differentiation was conducted by running a SIMPER test using the function 'simper()' with the 'vegan' R package (version 2.5-7).

3. Results

3.1. Pedoclimatic Conditions

Results of physical and chemical analyses of the soil are reported in Table S2. Based on the USDA soil textural triangle classification system [41], all the collected soil belongs to the silty loam class, except Gavigno 2 which is classified as loam. The soils of Gavigno, being located in a wooded area, are richer in organic matter and have a more acid pH than the soils of the Val di Chiana (hilly area) and Prato (urban area). The soils mostly differed in P_2O_5, K_2O, Mg, and Ca contents, and all samples showed different groupings when subjected to a principal component analysis (PCA) performed on the soil physicochemical parameters (Figure 1).

Climatic conditions are summarized in Figure S1. Significant variability between the timepoints (in months) and sampling sites was observed for both temperature and precipitation. The monthly GDD varied from 172.5 (April) to 511.15 (August), 219.3 to 606.25, and 327.35 to 694.2 for Gavigno, Val di Chiana, and Prato, respectively. As expected, Prato was the warmest site, with 2686.25 GDD, followed by Val di Chiana (2211.65 GDD) and Gavigno (1806.4 GDD). Gavigno was the wettest sampling site, reaching a cumulative precipitation level of 499 mm for the entire growing season, with maximum precipitation (149.2 mm) recorded in June and minimum precipitation (43.0 mm) recorded in July. The same trend was also observed in Val di Chiana, with maximum (105.2 mm) and minimum (14.2 mm) rainfalls recorded in June and July, respectively. Prato was the driest site, reaching a cumulative precipitation level of 239.2 mm from April to August, which is less than half of the levels found in Gavigno. Unexpectedly, August was the wettest month (70.2 mm), while July was confirmed as the least rainy month (22 mm).

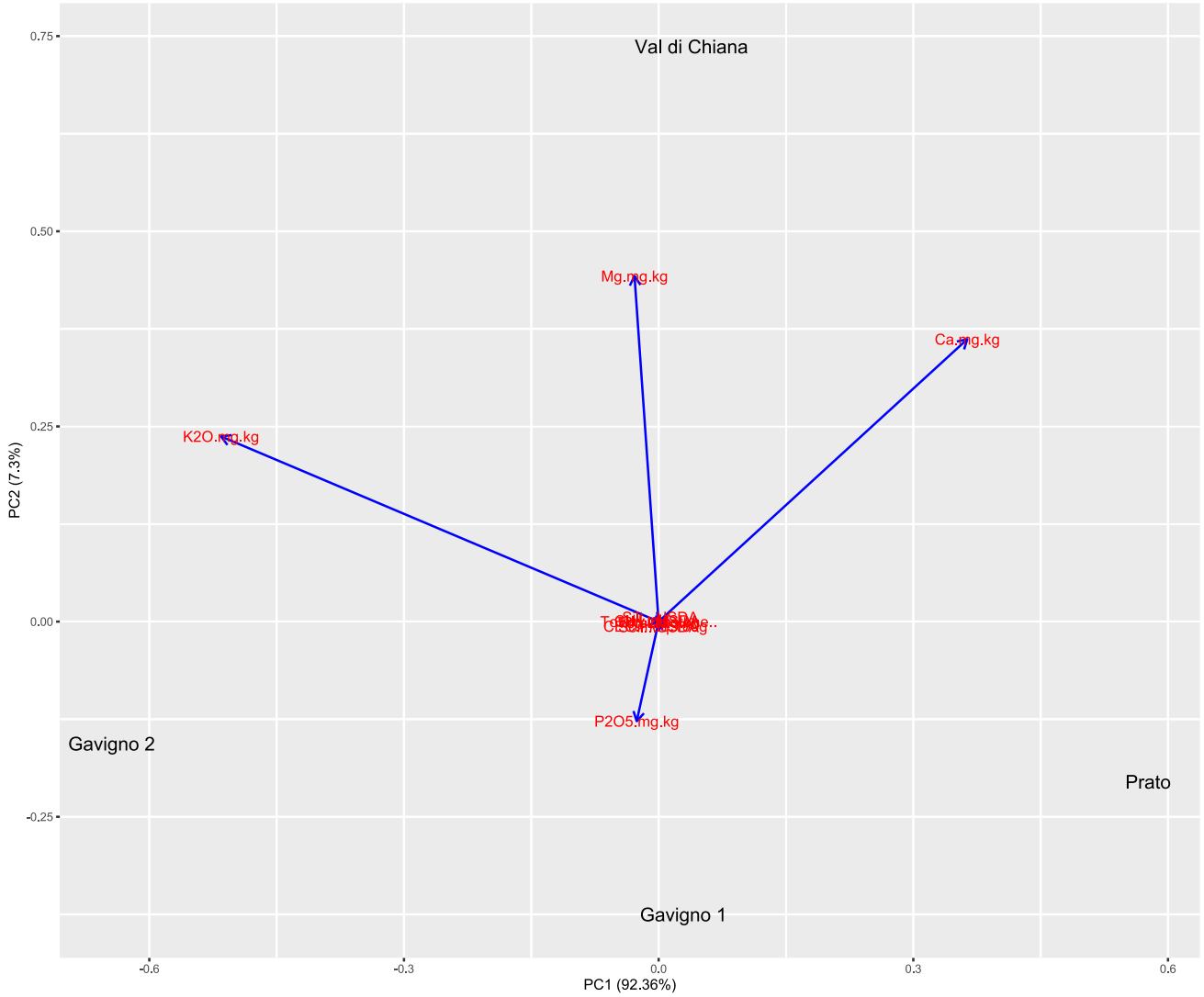


Figure 1. Influence of the physicochemical composition of sampled localities. A principal component analysis, with a biplot, is shown. The percentage of the variance of the first two components is reported. Overlapped labels on the biplot indicate that they do not contribute to the differences among localities.

3.2. Overall Diversity of the Rhizosphere Bacterial Microbiota

A total of 3′760′646 16S rRNA reads were obtained, and 2′721′248 (72% of total reads) passed quality filtering (Table S1). After the clustering step, a total of 14′332 ASVs were obtained (Dataset S1). Rarefaction curves obtained from the ASVs reached a plateau for all samples (Figure S2), indicating a satisfactory survey of the bacterial diversity (Good's coverage, Table S3).

3.3. Location Is the Main Driver of the Taxonomic Composition of Potato Bacterial Rhizomicrobiota

Alpha diversity estimates were obtained, and comparisons between location and phenological stages were performed. Wilcoxon tests indicated that the Simpson index and both evenness indices showed different values among localities (Figure 2, Table S4). Moreover, the PERMANOVA reported that the phenological stage, and the interaction between locality and stage, significantly contributed to samples differences (Tables 2 and S5).

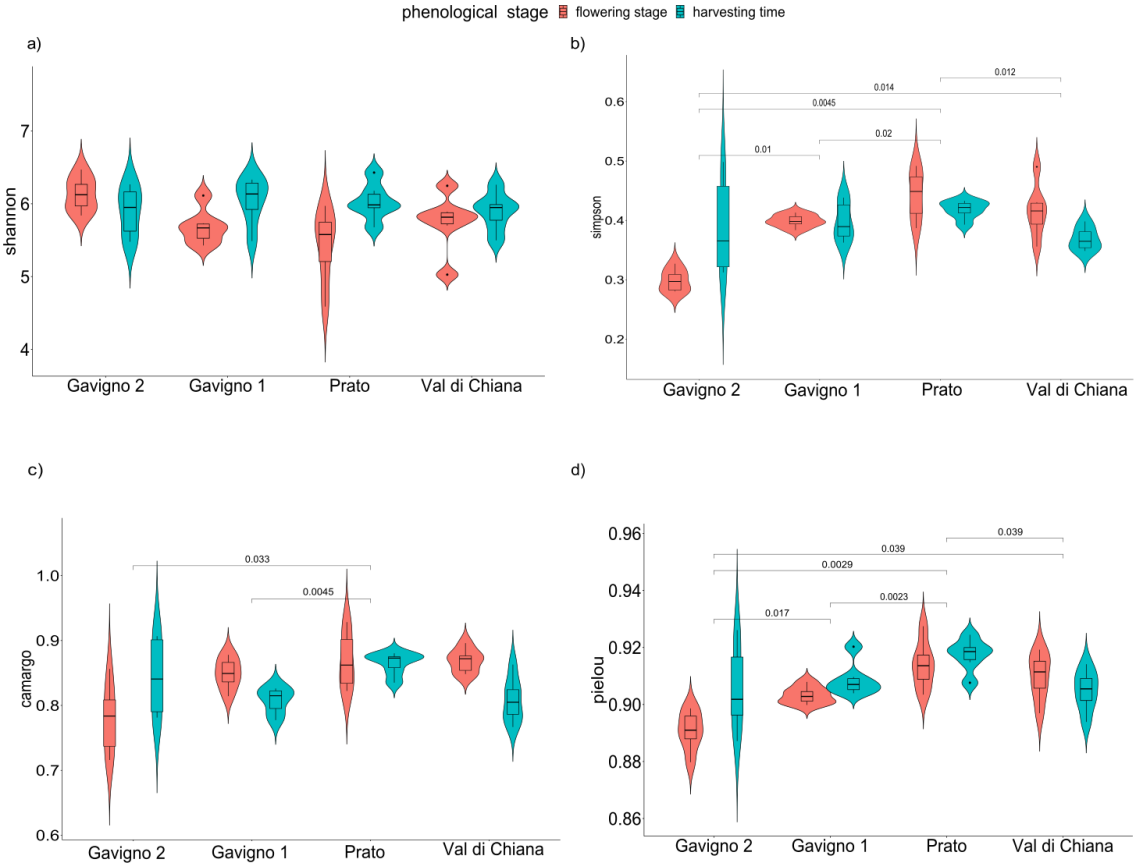


Figure 2. Alpha diversity indices of the samples. (**a**) Shannon index; (**b**) Simpson index; (**c**,**d**) evenness (Camargo's and Pielou's indices, respectively). Lines and the numbers above the lines indicate the *p*-values of the contrasts (Wilcoxon test).

Table 2. Results of PERMANOVAs on alpha diversity indices, showing the contributions of the locality, the phenological stage, and their interaction. Pseudo F-statistic *p*-values are reported. n.s., not significant ($p > 0.05$).

	Shannon		**Simpson**		**Pielou**		**Camargo**	
	F	***p-Value***	**F**	***p-Value***	**F**	***p-Value***	**F**	***p-Value***
Locality	15.064	n.s.	11.9274	0.001	6.2539	0.004	4.4182	0.011
Stage	52.131	0.018	0.4555	n.s.	2.7498	n.s.	0.6968	n.s.
Locality:stage	32.784	0.03	7.5625	0.002	4.7729	0.008	6.7443	0.001

The same pattern was observed for the taxonomic similarities among samples (Figure 3), with the two Gavigno sets of samples (Gavigno 1 and Gavigno 2) grouping together and remaining separated from the Prato and Val di Chiana samples. Within the same locality, the two phenological stages showed a separation. A PERMANOVA test (Table S6) confirmed the high relevance of localities (sites, $F = 1.64$, $p < 0.001$) but also confirmed the contribution of the phenological stage ($F = 1.17$, $p < 0.01$) and their interaction ($F = 1.19$, $p < 0.001$) in the sample differences.

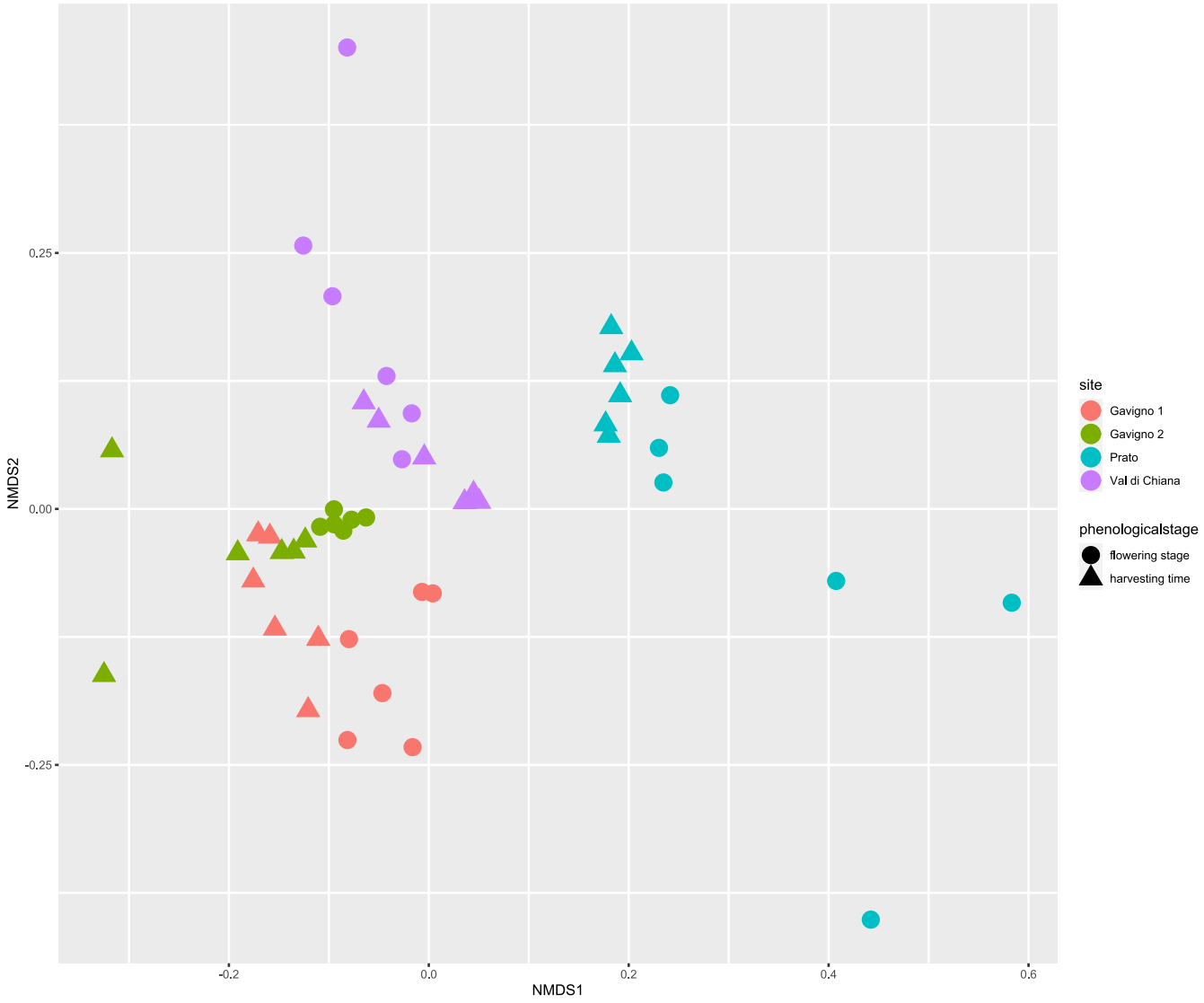


Figure 3. Nonmetric multidimensional scaling of potato bacterial rhizomicrobiota. Colors indicate locations; circles, flowering stages; triangles, harvesting stages. Stress value = 0.17.

3.4. Taxonomic Signatures of Localities and Phenological Stages

Figure 4 reports the relative abundances of taxonomies at the phylum level. The analysis of the taxonomic composition revealed that the most abundant phyla were Acidobacteriota, Actinobacteriota, Chloroflexi, and Proteobacteria.

In order to check for the presence of the differential abundance of taxonomies with respect to localities and phenological stages, the package DESeq2 was used in the overall dataset. While no single ASV was found to be statistically different in abundance, with respect to the phenological stage for all samples, 15 ASVs were found to be differentially distributed among localities—9 of which were affiliated with the class Actinobacteria (Table S7), suggesting the presence of soil-/locality-related microbiota signatures.

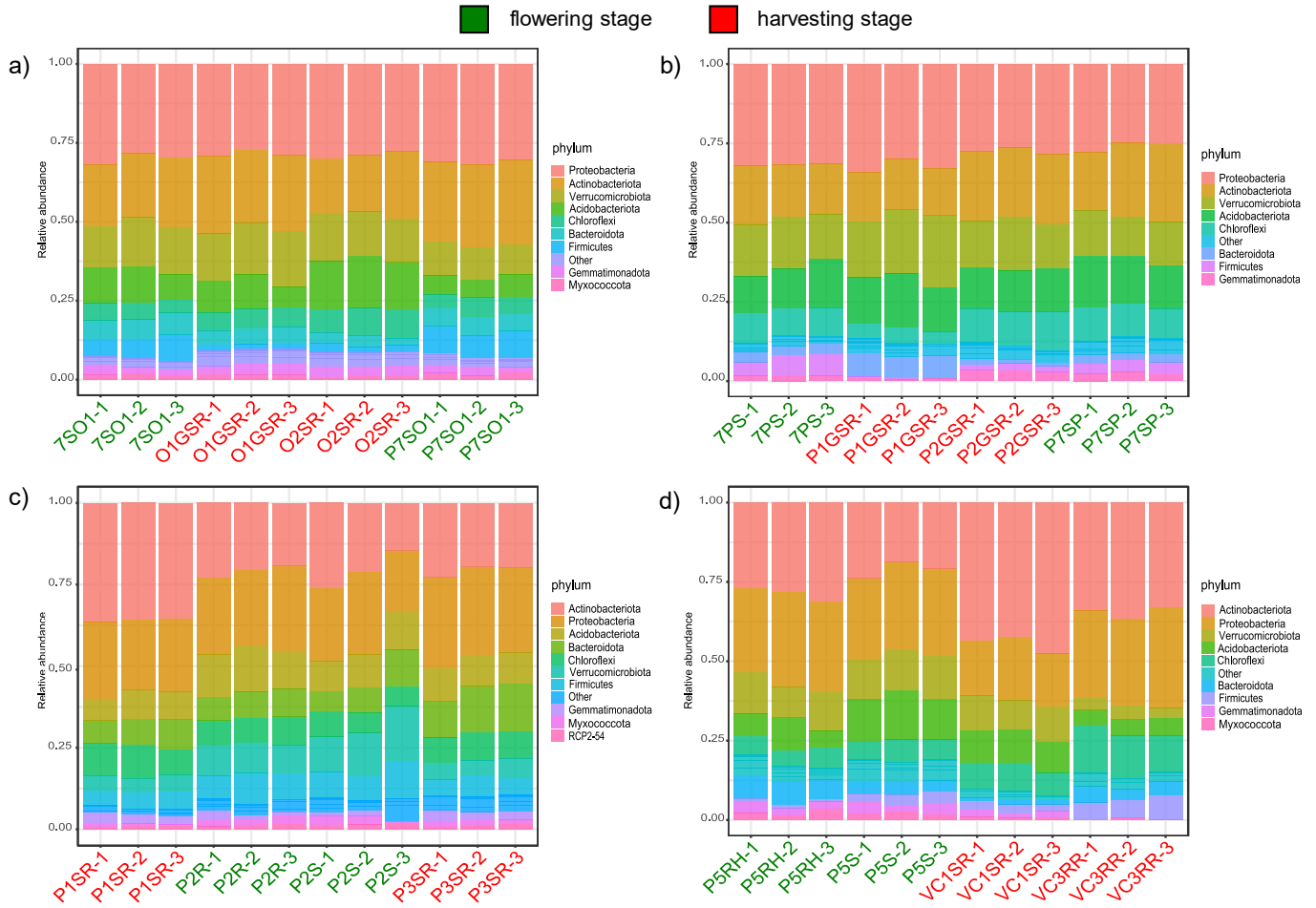


Figure 4. Relative abundances of bacterial phyla in the potato bacterial rhizomicrobiota. The four localities are shown separately. Samples from flowering and harvesting stages are indicated. See Table S1 for sample codes. (**a**) Gavigno 1. (**b**) Gavigno 2. (**c**) Prato. (**d**) Val di Chiana.

3.5. Functional Potential of the Potato Rhizomicrobiota

Functional profiles were inferred from retrieved taxonomies, in terms of the metabolic pathways present in the bacterial taxa of the overall potato rhizomicrobiota. We identified a total of 434 pathways distributed in all samples (Table S7). Using a SIMPER test, we identified nine and three pathways that significantly differentiated samples with respect to locality and phenological stage, respectively (Figure 5). The pathways included abilities related to the biosynthesis of molecules (ENTBACSYN-PWY, enterobactin biosynthesis), the degradation of plant metabolites (PWY-5266, p-cymene degradation; PWY-5273, p-cumate degradation), anaerobic metabolism (PWY0-1533, methylphosphonate degradation; PWY-7007, methyl ketone biosynthesis; PWY-6876, isopropanol biosynthesis), carbon fixation (PWY-5392, reductive TCA cycle), glycogen degradation (PWY-5941), pyrimidine biosynthesis (PWY-7187), and vitamin B12-related metabolism (adenosylcobalamin salvage from cobinamide II, PWY-6269). The latter was only found in relation to phenological stages differences.

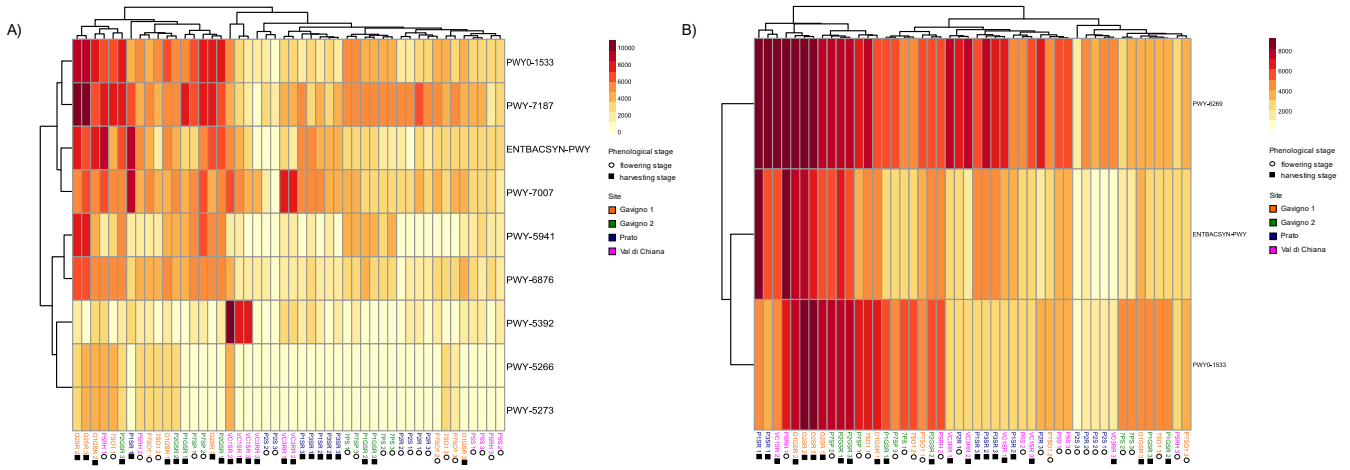


Figure 5. Metabolic pathways mainly contributing to differences among localities (**A**) and phenological stages (**B**). See Table S1 for samples codes.

4. Discussion

The main aim of this study was to identify how much the rhizosphere microbiota of potato plants is impacted by the site of cultivation and to examine whether the plant growth stage determines changes in the rhizomicrobiota. Though we are aware of the limitation of a study conducted during only one year and with a limited number of sites, the obtained results clearly indicated that the overall diversity of the microbiota was principally affected by the pedoclimatic conditions of the site of cultivation. However, the plant phenological stage was also shown to play a role in rhizomicrobiota diversity and taxonomic composition. Indeed, the plant developmental stage is a relevant factor influencing rhizosphere microbiota [8]. In a previous study [9], conducted with low-input agriculture in the center of the origin area of potatoes (High Andes), alpha diversity was influenced by the sampling site, while the effects of plant development on alpha diversity were less pronounced. This result is in agreement with evidence from our study, which was conducted using the same variety for all samples (cv. Kennebec). In fact, in our case-study also reported that alpha diversity mostly differs in relation to the locality, and, to a lesser extent, in relation to plant phenological stages. Interestingly, in the same study [9], taxonomic diversity was significantly affected by the phenological stage, which perfectly matches our results, indicating a stronger effect of the phenological stage on the taxonomic structure of the microbiota than that observed for alpha diversity. However, in the case of taxonomic diversity, the main effect was still related to the locality, justifying the possibility of identifying microbiota-related markers for the site of cultivation, as previously observed for other plant species, such as barley, *Crocus*, and orchids [16–18].

An abundant core microbiota, constituted by members of phyla Acidobacteriota, Actinobacteriota, Chloroflexi, and Proteobacteria, was found, irrespective of the cultivation site and phenological stage. Acidobacteria, Actinobacteriota, and Proteobacteria were found to be abundant in potatoes in other studies [9,10,42], indicating that differences in varieties, geographical locations, agronomic practices, and laboratory techniques (e.g., DNA extraction and sequencing methods) do not affect the overall taxonomic representation of bacterial potato rhizomicrobiota. Within these main phyla, Actinobacteriota were particularly interesting, since the ASVs related to the among-locality differences were taxonomically affiliated with this group. The presence of bacteria from the phylum Actinobacteriota (mainly from the genus *Streptomyces*) within these ASVs may suggest that the production of secondary metabolites, including novel antibiotics, could be an important function of the potato rhizomicrobiota. In fact, though a species from the genus *Streptomyces*, *S. scabiei*, is the etiological agent of potato common scab disease [43], other strains from the genus *Streptomyces* have shown important antagonistic activities and the suppression of common scab disease [44]. We can envisage a future investigation on *Streptomyces* associated with

potato rhizosphere and the isolation of strains with relevant features within this context. Additionally, the presence of ASVs differing by locality also implies the possibility of developing DNA-based markers for such strains to allow for the identification of cultivation sites and the protection of local productions, which is a relevant aspect for traditional Italian food chains [22]. Finally, it is worth noting that the most abundant pathways represented after the PiCRUST analysis of the microbiota were those related to enterobactin biosynthesis and the utilization of phenolic metabolites (p-cymene, p-cumate). Enterobactin is a well known siderophore, helping cells to chelate iron ions and transport them into the cell [45]. Siderophore production is a key function of plant-associated bacteria [46,47] and, in particular, of growth-promoting bacteria [48]. This is highly relevant for bioinoculant formulations, where strains with important, plant growth-promoting functions are combined. Competition for iron is also relevant for bacterial potato pathogens, such as *Dickeya* spp. [49], and a relevant interplay between the plant and the microbiota exists for iron scavenging [50]. We can, consequently, look forward to the isolation of strains with such relevant functions from the potato rhizomicrobiota as ingredients for future bioinoculant formulations to help potatoes with microelement nutrition [51]. The presence of abilities to degrade terpenes is also relevant and highlights the metabolic interaction between the host plant and the microbiota, as observed in other systems [52–55]. Potatoes—and cv. Kennebec, in particular—are rich in terpene compounds [56], which are present on the skin of tubers and released during growth by the root as the root exudates.

In conclusion, this work reinforces the importance of the site of cultivation for the potato microbiota.It may also suggest the presence of strains and functions that can be further inspected for their possible use in sustainably improving potato cultivation through bioinoculants and developing microbiota-based fingerprinting of potatoes. However, we should be cautious in interpreting present results from a generalized view. In fact, it would be interesting to know how stable this pattern of rhizomicrobiota is in different varieties and throughout the years, especially when farmers change fertilization methods and when different temperatures and different amounts of precipitation occur. Climatic conditions can influence bacterial and fungal pathogens, which, in turn, can influence the overall rhizomicrobiota. Indeed, since sites differed by climatic conditions, additional work should be conducted to evaluate relationships between climatic conditions, rhizomicrobiota, and plant productivity at this smaller geographical scale to better integrate microbiota variation in an agronomic context.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/microorganisms10091743/s1. Dataset S1: Table reporting the amplicon sequence variants (ASVs) detected, their taxonomic assignment and occurrence. Table S1: Table containing, sample codes, locality, phenological stage, fast1 file name, the number of total input reads from fastq files and the number of reads that passed the filtering parameters. Table S2: Table containing the physicochemical features of soil samples. Table S3: Table of Good's coverage values. Good's coverage was estimated on the amplicon sequence variant (ASV) table for each sample after removing chimeras. Table S4: List of richness, diversity and evenness indices; Table S5: Results of Permutational Multivariate Analysis of Variance (PERMANOVA) test performed on alpha diversity indicates. Table S6: List of p values calculated through a Permutational Multivariate Analysis of Variance (PERMANOVA) test performed on the phyloseq object after a centered log ratio (CLR) transformation. Table S7: Results of DeSeq2 on ASVs mostly differing among localities. Table S8: List of all identified metabolic pathways present in the taxa retrieved in the bacterial community. Figure S1: Climatic conditions. The cumulated growing degree days (GDD; °C) (top row) and the cumulated precipitation (mm) (bottom row) are reported. Results were calculated on a monthly basis (April-August) (left panel) and for the entire growing season (from sowing to maturity) (right panel). Figure S2: Rarefaction curve of all soil samples. The x axis indicates the number of sequences found for each sample, while the y axis indicates the species richness.

Author Contributions: Conceptualization, M.M., A.B., S.O., A.M., and L.C.; methodology, L.C., A.B.; validation, A.B., M.M., L.C., C.F., and F.V.; formal analysis, L.C.; investigation, L.C., C.F., F.V., and A.B.; resources, A.B., M.M., S.O., and A.M.; data curation, A.M., L.C.; writing—original draft preparation, A.M., L.C.; writing—review and editing, A.B., M.M., S.O., A.M., L.C., and C.F.; supervision, A.M., and L.C.; project administration, A.M.; funding acquisition, A.M., S.O., and M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by the grant MICRO4Legumes, D.M.n.89267 (Italian Ministry of Agriculture) provided to AM. LC is supported by a PhD fellowship from MICRO4Legumes.

Data Availability Statement: The sequences dataset was deposited in the SRA database under the BioProject PRJNA730792.

Acknowledgments: We express our gratitude to the farmers who kindly granted permission to sample their production fields, Az. Agr. Colzi Paolo (Prato, Italy) and Az. Agr. Fonte dei Piani (Gavigno, Italy). We warmly acknowledge Ewa Lojkowska (University of Gdansk, Poland) for the critical reading of the manuscript and fruitful discussion on potato microbiota.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Motyka-Pomagruk, A.; Zoledowska, S.; Sledz, W.; Lojkowska, E. The occurrence of bacteria from different species of Pectobacteriaceae on seed potato plantations in Poland. *Eur. J. Plant Pathol.* **2021**, *159*, 309–325. [CrossRef]
2. Busby, P.E.; Soman, C.; Wagner, M.R.; Friesen, M.L.; Kremer, J.; Bennett, A.; Morsy, M.; Eisen, J.A.; Leach, J.E.; Dangl, J.L. Research priorities for harnessing plant microbiomes in sustainable agriculture. *PLoS Biol.* **2017**, *15*, e2001793. [CrossRef]
3. Schlaeppi, K.; Bulgarelli, D. The Plant Microbiome at Work. *Mol. Plant-Microbe Interact. MPMI* **2015**, *212*, 212–217. [CrossRef]
4. Compant, S.; Samad, A.; Faist, H.; Sessitsch, A. A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. *J. Adv. Res.* **2019**, *19*, 29–37. [CrossRef]
5. Sánchez-Cañizares, C.; Jorrín, B.; Poole, P.S.; Tkacz, A. Understanding the holobiont: The interdependence of plants and their microbiome. *Curr. Opin. Microbiol.* **2017**, *38*, 188–196. [CrossRef]
6. Vandenkoornhuyse, P.; Quaiser, A.; Duhamel, M.; Le Van, A.; Dufresne, A. The importance of the microbiome of the plant holobiont. *New Phytol.* **2015**, *206*, 1196–1206. [CrossRef]
7. Agler, M.T.; Ruhe, J.; Kroll, S.; Morhenn, C.; Kim, S.T.; Weigel, D.; Kemen, E.M. Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. *PLoS Biol.* **2016**, *14*, e1002352. [CrossRef]
8. Chaparro, J.M.; Badri, D.V.; Vivanco, J.M. Rhizosphere microbiome assemblage is affected by plant development. *ISME J.* **2014**, *8*, 790–803. [CrossRef]
9. Pfeiffer, S.; Mitter, B.; Oswald, A.; Schloter-Hai, B.; Schloter, M.; Declerck, S.; Sessitsch, A. Rhizosphere microbiomes of potato cultivated in the High Andes show stable and dynamic core microbiomes with different responses to plant development. *FEMS Microbiol. Ecol.* **2017**, *93*, fiw242. [CrossRef]
10. Hou, Q.; Wang, W.; Yang, Y.; Hu, J.; Bian, C.; Jin, L.; Li, G.; Xiong, X. Rhizosphere microbial diversity and community dynamics during potato cultivation. *Eur. J. Soil Biol.* **2020**, *98*, 103176. [CrossRef]
11. Zimudzi, J.; van der Waals, J.E.; Coutinho, T.A.; Cowan, D.A.; Valverde, A. Temporal shifts of fungal communities in the rhizosphere and on tubers in potato fields. *Fungal Biol.* **2018**, *122*, 928–934. [CrossRef]
12. Inceoğlu, Ö.; Salles, J.F.; Van Overbeek, L.; Van Elsas, J.D. Effects of plant genotype and growth stage on the betaproteobacterial communities associated with different potato cultivars in two fields. *Appl. Environ. Microbiol.* **2010**, *76*, 3675–3684. [CrossRef]
13. Riva, V.; Mapelli, F.; Bagnasco, A.; Mengoni, A.; Borin, S.S. A Meta-Analysis Approach to De fi ning the Culturable Core of Plant Endophytic Bacterial Communities. *Appl. Environ. Microbiol.* **2022**, *88*, e02537-21. [CrossRef]
14. Cangioli, L.; Vaccaro, F.; Fini, M.; Mengoni, A. Scent of a Symbiont: The Personalized Genetic Relationships of Rhizobium—Plant Interaction. *Int. J. Mol. Sci.* **2022**, *23*, 3358. [CrossRef]
15. Vaccaro, F.; Cangioli, L.; Mengoni, A.; Fagorzi, C. Synthetic plant microbiota challenges in nonmodel species. *Trends Microbiol.* **2022**. [CrossRef]
16. Alegria Terrazas, R.; Balbirnie-Cumming, K.; Morris, J.; Hedley, P.E.; Russell, J.; Paterson, E.; Baggs, E.M.; Fridman, E.; Bulgarelli, D. A footprint of plant eco-geographic adaptation on the composition of the barley rhizosphere bacterial microbiota. *Sci. Rep.* **2020**, *10*, 12916. [CrossRef]
17. Lin, M.; Xiong, H.; Xiang, X.; Zhou, Z.; Liang, L.; Mei, Z. The effect of plant geographical location and developmental stage on root-associated microbiomes of Gymnadenia conopsea. *Front. Microbiol.* **2020**, *11*, 1257. [CrossRef]
18. Bhagat, N.; Sharma, S.; Ambardar, S.; Raj, S.; Trakroo, D.; Horacek, M.; Zouagui, R.; Sbabou, L.; Vakhlu, J. Microbiome Fingerprint as Biomarker for Geographical Origin and Heredity in Crocus sativus: A Feasibility Study. *Front. Sustain. Food Syst.* **2021**, *5*, 688393. [CrossRef]
19. FAOSTAT. *Primary Crops Statistical Database*; FAO: Rome, Italy, 2004.

20. Menna, C.; Tosco, D.; Colantuoni, F.; Caracciolo, F.; sannino, G.; Cicia, G.; Cembalo, L.; D'amico, M.; Lombardi, P.; Maraglino, T. Country-of-origin labelling for the Italian early potato supply chain. *New Medit Mediterr. J. Econ. Agric. Environ. Rev. Méditerranéenne d'Economie Agric. Environ.* **2013**, *12*, 37.
21. Ingallina, C.; Spano, M.; Sobolev, A.P.; Esposito, C.; Santarcangelo, C.; Baldi, A.; Daglia, M.; Mannina, L. Characterization of Local Products for Their Industrial Use: The Case of Italian Potato Cultivars Analyzed by Untargeted and Targeted Methodologies. *Foods* **2020**, *9*, 1216. [CrossRef]
22. Manzelli, M.; Romagnoli, S.; Ghiselli, L.; Benedettelli, S.; Palchetti, E.; Andrenelli, L.; Vecchio, V. Typicity in potato: Characterization of geographic origin. *Ital. J. Agron.* **2010**, *5*, 61–68. [CrossRef]
23. Pacifico, D. Upland Italian Potato Quality—A Perspective. *Sustainability* **2018**, *10*, 3939. [CrossRef]
24. Arnold, C.Y. The determination and significance of the base temperature in a linear heat unit system. *Proc. Am. Soc. Hortic. Sci.* **1959**, *74*, 430–445.
25. Mix, K.; Rast, W.; Lopes, V.L. Increases in growing degree days in the Alpine Desert of the San Luis Valley, Colorado. *Water. Air. Soil Pollut.* **2010**, *205*, 289–304. [CrossRef]
26. Singh, B.K.; Trivedi, P.; Egidi, E.; Macdonald, C.A.; Delgado-Baquerizo, M. Crop microbiome and sustainable agriculture. *Nat. Rev. Microbiol.* **2020**, *18*, 601–602. [CrossRef]
27. Cangioli, L.; Mancini, M.; Napoli, M.; Fagorzi, C.; Orlandini, S.; Vaccaro, F. Differential Response of Wheat Rhizosphere Bacterial Community to Plant Variety and Fertilization. *Int. J. Mol. Sci.* **2022**, *23*, 3616. [CrossRef]
28. Klindworth, A.; Pruesse, E.; Schweer, T.; Peplies, J.; Quast, C.; Horn, M.; Glöckner, F.O. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **2013**, *41*, e1. [CrossRef]
29. Checcucci, A.; Maida, I.; Bacci, G.; Ninno, C.; Bilia, A.R.; Biffi, S.; Firenzuoli, F.; Flamini, G.; Fani, R.; Mengoni, A. Is the plant-associated microbiota of *Thymus* spp. adapted to plant essential oil? *Res. Microbiol.* **2017**, *168*, 276–282. [CrossRef]
30. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13*, 581–583. [CrossRef]
31. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, D590–D596. [CrossRef]
32. Wright, E.S.; Yilmaz, L.S.; Noguera, D.R. DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. *Appl. Environ. Microbiol.* **2012**, *78*, 717–725. [CrossRef]
33. McMurdie, P.J.; Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **2013**, *8*, e61217. [CrossRef]
34. Shetty, S.A.; Lahti, L. Microbiome data science. *J. Biosci.* **2019**, *44*, 115. [CrossRef]
35. Camargo, J.A. On Measuring Species Evenness and Other Associated Parameters of Community Structure. *Oikos* **1995**, *74*, 538–542. [CrossRef]
36. Pielou, E.C. The measurement of diversity in different types of biological collections. *J. Theor. Biol.* **1966**, *13*, 131–144. [CrossRef]
37. Wickham Hadley. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2009; ISBN 978-0-387-98140-6.
38. Kandlikar, G.S.; Gold, Z.J.; Cowen, M.C.; Meyer, R.S.; Freise, A.C.; Kraft, N.J.B.; Moberg-Parker, J.; Sprague, J.; Kushner, D.J.; Curd, E.E. ranacapa: An R package and Shiny web app to explore environmental DNA data with exploratory statistics and interactive visualizations. *F1000Research* **2018**, *7*, 1734. [CrossRef]
39. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef]
40. Feng, K.; Zhang, Z.; Cai, W.; Liu, W.; Xu, M.; Yin, H.; Wang, A.; He, Z.; Deng, Y. Biodiversity and species competition regulate the resilience of microbial biofilm community. *Mol. Ecol.* **2017**, *26*, 6170–6182. [CrossRef]
41. Staff, S.S.L. *Soil Survey Laboratory Methods Manual*; National Soil Survey Center: Lincoln, NE, USA, 1996.
42. Buchholz, F.; Antonielli, L.; Kostić, T.; Sessitsch, A.; Mitter, B. The bacterial community in potato is recruited from soil and partly inherited across generations. *PLoS ONE* **2019**, *14*, e0223691. [CrossRef]
43. Liu, J.; Nothias, L.-F.; Dorrestein, P.C.; Tahlan, K.; Bignell, D.R.D. Genomic and Metabolomic Analysis of the Potato Common Scab Pathogen Streptomyces scabiei. *ACS Omega* **2021**, *6*, 11474–11487. [CrossRef]
44. Hiltunen, L.H.; Tarvainen, O.; Kelloniemi, J.; Tanskanen, J.; Karhu, J.; Valkonen, J.P.T. Soil bacterial community in potato tuberosphere following repeated applications of a common scab suppressive antagonist. *Appl. Soil Ecol.* **2021**, *167*, 104096. [CrossRef]
45. Qi, B.; Han, M. Microbial Siderophore Enterobactin Promotes Mitochondrial Iron Uptake and Development of the Host via Interaction with ATP Synthase. *Cell* **2018**, *175*, 571–582.e11. [CrossRef]
46. Ling-yun, H.; Kyle, W.D.; Helene, A.-P.; Michael, M.; Barak, J.D. Requirement of Siderophore Biosynthesis for Plant Colonization by Salmonella enterica. *Appl. Environ. Microbiol.* **2012**, *78*, 4561–4570. [CrossRef]
47. Vandana, U.K.; Rajkumari, J.; Singha, L.P.; Satish, L.; Alavilli, H.; Sudheer, P.D.V.N.; Chauhan, S.; Ratnala, R.; Satturu, V.; Mazumder, P.B.; et al. The Endophytic Microbiome as a Hotspot of Synergistic Interactions, with Prospects of Plant Growth Promotion. *Biology* **2021**, *10*, 101. [CrossRef]
48. Oleńska, E.; Małek, W.; Wójcik, M.; Swiecicka, I.; Thijs, S.; Vangronsveld, J. Beneficial features of plant growth-promoting rhizobacteria for improving plant growth and health in challenging conditions: A methodical review. *Sci. Total Environ.* **2020**, *743*, 140682. [CrossRef]

49. Perkowska, I.; Potrykus, M.; Siwinska, J.; Siudem, D.; Lojkowska, E.; Ihnatowicz, A. Interplay between coumarin accumulation, iron deficiency and plant resistance to *Dickeya* spp. *Int. J. Mol. Sci.* **2021**, *22*, 6449. [CrossRef]
50. Stassen, M.J.J.; Hsu, S.H.; Pieterse, C.M.J.; Stringlis, I.A. Coumarin Communication Along the Microbiome–Root–Shoot Axis. *Trends Plant Sci.* **2021**, *26*, 169–183. [CrossRef]
51. Pattnaik, S.; Mohapatra, B.; Gupta, A. Plant Growth-Promoting Microbe Mediated Uptake of Essential Nutrients (Fe, P, K) for Crop Stress Management: Microbe–Soil–Plant Continuum. *Front. Agron.* **2021**, *58*, 689972. [CrossRef]
52. Xiong, C.; Zhu, Y.; Wang, J.; Singh, B.; Han, L.; Shen, J.; Li, P.; Wang, G.; Wu, C.; Ge, A.; et al. Host selection shapes crop microbiome assembly and network complexity. *New Phytol.* **2020**, *229*, 1091–1104. [CrossRef]
53. Trivedi, P.; Leach, J.E.; Tringe, S.G.; Sa, T.; Singh, B.K. Plant–microbiome interactions: From community assembly to plant health. *Nat. Rev. Microbiol.* **2020**, *18*, 607–621. [CrossRef]
54. Abedini, D.; Jaupitre, S.; Bouwmeester, H.; Dong, L. Metabolic interactions in beneficial microbe recruitment by plants. *Curr. Opin. Biotechnol.* **2021**, *70*, 241–247. [CrossRef]
55. diCenzo, G.C.; Tesi, M.; Pfau, T.; Mengoni, A.; Fondi, M. Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. *Nat. Commun.* **2020**, *11*, 2574. [CrossRef]
56. Giupponi, L.; Leoni, V.; Pedrali, D.; Ceciliani, G.; Bassoli, A.; Borgonovo, G. Morphometric and Phytochemical Characterization and Elevation Effect on Yield of Three Potato Landraces of the Ligurian Apennines (Northern Italy). *J. Appl. Bot. Food Qual.* **2020**, *93*, 234–243. [CrossRef]

CellPress

Trends in Microbiology

Forum

Synthetic plant microbiota challenges in nonmodel species

Francesca Vaccaro,[1] Lisa Cangioli,[1] Alessio Mengoni,[1,*] and Camilla Fagorzi[1]

Plant-associated microbiota are becoming central in the development of ways to improve plant productivity and health. However, most research has focussed mainly on a few model plant species. It is essential to translate discoveries to the many nonmodel crops, allowing the design and application of effective synthetic microbiota.

Opportunities from synthetic microbial communities in sustainable agriculture

Recent studies estimate that the world population will reach 9.8 billion people by 2050. To meet the increasing demand for food, feed, and bioenergy, agricultural production should increase by 70% but the use of artificial fertilizers, such as synthetic nitrogen, is energy-intensive and is not environmentally sustainable [1]. A very innovative and ecofriendly strategy to face the agricultural challenges is represented by the naturally occurring microbial systems [2]. Microbial communities are ubiquitous in the environment and they have a pivotal role in the production of foods, the recycling of micronutrients, and in maintaining the health of humans, animals, and plants. Plant-associated microbes represent an enormous potential because of their plant-growth-promoting (PGP) traits. Thanks to the recent advances in 'omics technologies, and to the innovative computational tools, it has been possible to systematically explore and discover novel gene functions in microbiomes and identify and characterize biosynthetic gene clusters to produce secondary metabolites [3]. The exploration of plant microbiomes is expected to increase the exploitation of these gene functions. Biocontrol, biofertilization, and biostimulation can now be engineered [2].

Conventional inoculants cannot establish durable and efficient associations with plants under field conditions, leading to unsatisfactory results [4]. A number of strategies, from seed coating to using carriers, are on the market, but in many cases they still rely on empirical evidence [5]. The assessment of synthetic microbial communities (i.e., phytomicrobiome engineering, when talking about crop productivity) allows us to overcome this limitation by allocating functional components between two or more subpopulations of cells by means of the rational engineering of living organisms. Additionally, such synthetic communities are also a tool for discovering plant–microbe interactions and to predict the host–plant phenotype.

Prioritizing nonmodel plants

Moving from laboratory-scale evidence of synthetic microbial communities to field application is one of the greatest challenges of microbiome studies, on both the animal (human) and plant levels [6]. Concerning the plant microbiome, model plant species have been deeply studied and an understanding of the functions and interaction between the plant and the microbes is becoming more and more clear [7]. However, some model species, such as *Arabidopsis*, can have limitations, due for instance to the short life span (i.e., which does not extend to an entire year or several years, possibly limiting the knowledge about seasonal dynamics of the microbiota) or to the lack of interactions with a relevant portion of the soil microbiota (e.g., arbuscular mycorrhizal fungi, with which Brassicaceae do not form symbiosis [8]). The work on cereals – such as maize, rice, wheat, and barley – has shown the impact of domestication on the root-associated microbial communities [9], indicating that varieties and wild nonmodel relatives offer the opportunity to investigate the evolutionary interactions between microbes and plants.

It is becoming clear that we still do not know much about the peculiarities of the plant microbiota with respect to plant species, especially those of relevant use in agriculture, but still neglected as models in laboratory research. In a tentative list of plant species, we may include crops, such as several forage and grain legumes (e.g., alfalfa, chickpea, lentil), landraces of cereals (e.g., rice, maize, and wheat), which can have relatives as model species (such as *Medicago truncatula* for alfalfa) but which hold their own specific features (such as peculiar structure and physiology). Relevant examples emerged a few years ago from landraces of maize [10], and from experiments with field varieties and nonlaboratory bacterial strains in forage legumes (*Medicago sativa*, alfalfa) [11] (Box 1). These examples reflect the relevance and opportunities for investigation of nonmodel species in plant–microbe interaction studies in terms of microbial biodiversity discovery and importance of metabolic and signaling complexity for plant colonization. Moreover, tetraploid alfalfa and local wheat germplasm may also offer the opportunity to investigate basic biology phenomena, such as the effect that polyploidization and hybridization have on the recruitment of plant-associated microbiota and on the plant holobiont phenotypic plasticity. Such knowledge can be key for plant-breeding programs, focused on novel varieties with improved yield, higher resilience, and less nutritional requirement than current ones.

Ideally, research priorities should be given to considering plant germplasm for which close relatives, such as model species, exist, in order to allow an easier translation

Box 1. Examples of microbiological studies on nonmodel plants

The extensive investigations on legume–rhizobia and cereal model systems have paved the way to the development of laboratory and computational methods for establishing mechanistic relationships between plant and microbiota. Such methods, once translated to nonmodel plant species, will allow us to reduce the gap between the laboratory and the field. Examples are studies on the model legume *M. truncatula*, easily translated to the crop alfalfa (*M. sativa*). Here, the use of local varieties shed light on the existence of genotypic interactions that influence the phenotypic traits of the symbiosis [11]. In particular, a large fraction of microbial gene expression in the early symbiotic steps was influenced by the genotype of the variety of host plant tested (16%), by the strain genotype (35%), and by strain-by-host plant genotype interactions (29%), indicating that alfalfa can largely benefit from breeding programs taking into account the selectivity of the symbiotic rhizobia recruitment.

Shifting from the single-strain interaction model to microbiomes, comparison studies between rhizospheric microbiomes of modern model crops, wild relatives and landraces of barley (*Hordeum vulgare*) represent an opportunity to investigate the structure and the evolution of the microbiota and are ultimately critical to design synthetic communities [9].

Because of their economic interest, diazotrophic microbial associations with cereals arouse great interest. In the case of indigenous maize landraces, biological nitrogen fixation is supported by a microbiota found in the mucilage secreted by an aerial root system. This example highlights the importance of research on nonmodel systems [10]. In this case, the fact that other cereal crops produce these secretions from underground and aerial roots could allow the transfer of this trait by breeding.

of methods and knowledge gained from such model species to the nonmodel germplasm. However, in a world increasingly more subjected to climate change, reliance on a limited number of crops raises questions on food and nutritional security. Therefore, the knowledge on a few model plants and crops needs to be translated to several nonmodel crops, which are generally more resilient to environmental stresses but which suffer from the lack of information on associated microbes, which can assist and promote yield, resistance, and nutrition in a sustainable agriculture scenario. Underutilized and climate-resilient crops – such as the Chenopodiaceae *Amaranthus* spp., and *Chenopodium quinoa* (quinoa) – but also cereals such as *Eragrostis tef* (teff) and *Setaria italica* (foxtail millet), are local sources of food, characterized by high biological value protein content (https://cropsforthefutureuk.org/), which can largely benefit from nitrogen-fixing synthetic microbiota applications to increase their yield and nutritional value with low-input agricultural management. Prioritizing microbiological investigations on such nonmodels, using the methods and knowledge learned from model species, is urgently needed for securing future food security.

Translating from models to nonmodels

Systems biology has emerged as an effective top-down approach for designing, predicting, and manipulating the behavior of microbial populations. This approach allows us to rationally assemble microbial consortia and gives the opportunity to provide an initial framework to predict how the application of synthetic microbial communities can affect nonmodel crop health and yield. Complex biological phenomena, such as the behavior of microbial consortia, can be simulated by (i) dynamic computational models that can predict how the community will respond, given a parameter, (ii) stoichiometric models, such as the flux balance analysis (FBA) that allows us to simulate metabolic networks, based on genome-scale reconstructions (GENRE) and stoichiometric coefficients of every single cellular reaction, or (iii) agent-based models to simulate the behaviors of consortia over time [12].

Bacterial genome-wide association analysis (GWAS) is another tool to define genetic determinants related to phenotypic traits of interest, and a model of epistasis can be applied to define higher order interactions among strains and between strains and host plant [13].

Once a hypothetical community is assembled, these modeling approaches are able to predict growth capability, intracellular reactions, metabolic phenotypes, competition or synergism, based on molecular crosstalk [4]. Moreover, not only is it possible to model community metabolic networks, but also to develop a holistic *in silico* representation of the integrated metabolism of the holobiont as proposed few years ago for *Medicago truncatula–Sinorhizobium meliloti* symbiosis [14]. Finally, the ideal synthetic microbial community, once inoculated in the soil, should grow and persist in the presence of the indigenous microbiome and local environmental conditions. As any community, it will be influenced by four high-level processes, namely (i) selection, (ii) dispersal, (iii) drift, and (iv) speciation. Integrating community ecology theory with systems biology into plant microbiome research is needed to develop a more holistic and mechanistic synthesis in plant microbiome research [15].

Such technical approaches and interpretative models are very relevant for translating knowledge gained on controlled (and often simplified) plant model systems to nonmodel species (Figure 1). Years ago research priorities were identified [16] for addressing key questions related to the improvement of yield and stress resistance in crops. The identification of core microbiota and the recognition of the cultivar-specific response to microbiota are key to the bioinoculant preparation, as well as a precise understanding of strain dynamics, compatibility, and persistence in the synthetic community. Here, a possible limiting factor can be found in the culturability of strains of the microbiota, which consequently affect their investigation and use. Strategies of culturomics, by growing microbes on plant-derived substrates, but also by growing as consortia 'per se'

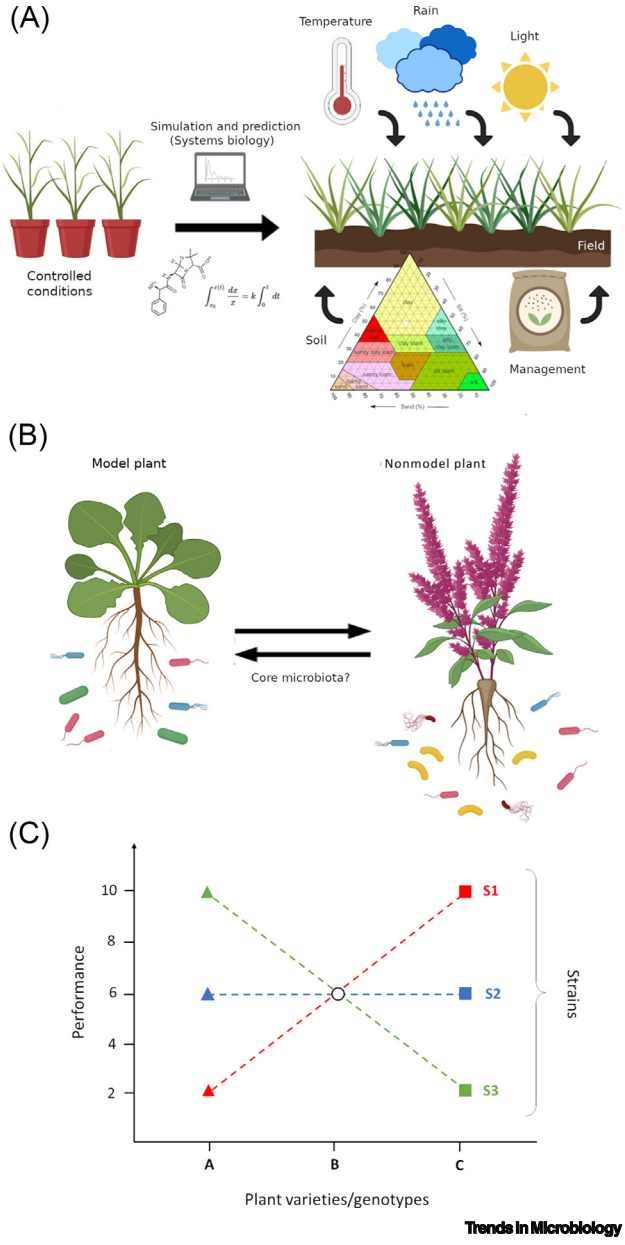


Figure 1. Key points and questions on translating knowledge of microbiota from model to nonmodel plants. (A) In the translation of acquired knowledge on the microbiota of model plants to nonmodel plants it is necessary to take into account the impact and synergy, over the bioinoculant, of environmental variables (a classical from-lab-to-the-field problem), including climatic conditions, soil composition, and crop agronomical management. Mathematical modeling and simulations based on systems biology analyses using multi-omics and synthetic microbiota data offer the opportunity to shorten the gap between laboratory studies on model plants to the assessment of microbial inoculant efficacy in the field. (B) To facilitate systems biology predictions, and allow to propose to the market a microbial inoculant with efficacy on a broad panel of plant species, the identification of a core plant microbiota, that is, a number of microbial taxa shared among multiple crops and between model and nonmodel species, needs a special effort. Moreover, studies on model plants have shown that plant performance is dependent on both additive and nonadditive interactions between plant genotype and microbial genotype. Studies deciphering such interactions (C), identifying strains which have a differential impact on plant performance in a genotype-dependent way, are compulsory. An example of genotype × genotype interaction is reported (S1, S3, plant genotype-dependent effect; S2, plant genotype-independent effect).

and not as pure cultures only, should be implemented to increase the culturable fraction of the microbiota. Finally, the effect of management and environments on plant–microbiome interactions still deserve increased attention. Here, the abovementioned systems biology methods can help translating from the finding on model species (where controlled experiments are available) to nonmodel species and field applications, where statistical methods allow a reduction in number and size of the field trials and may result in a faster delivery to the market.

Acknowledgment

This work was supported by the grant MICRO4Legumes, D.M. n. 89267 (Italian Ministry of Agriculture) to A.M. L.C. is supported by a PhD fellowship from MICRO4Legumes. F.V. is supported by a PhD fellowship from PON Research and Innovation 2014–2020, D.M. n. 1061/2021 (Italian Ministry of University and Research). C.F. is supported by a post-doctoral fellowship from the H2020 ERA-NETs SUSFOOD2 and CORE Organic Cofund, under the Joint SUSFOOD2/CORE Organic Call 2019.

Declaration of interests

No interests are declared.

[1]Department of Biology, University of Florence, Florence, Italy

*Correspondence:
alessio.mengoni@unifi.it (A. Mengoni).

https://doi.org/10.1016/j.tim.2022.06.006

© 2022 Elsevier Ltd. All rights reserved.

References

1. Le Mouël, C. and Forslund, A. (2017) How can we feed the world in 2050? A review of the responses from global scenario studies. *Eur. Rev. Agric. Econ.* 44, 541–591
2. Ke, J. *et al.* (2021) Microbiome engineering: synthetic biology of plant-associated microbiomes in sustainable agriculture. *Trends Biotechnol.* 39, 244–261
3. Levy, A. *et al.* (2018) Genomic features of bacterial adaptation to plants. *Nat. Genet.* 50, 138–150
4. de Souza, R.S.C. *et al.* (2020) From microbiome to traits: designing synthetic microbial communities for improved crop resiliency. *Front. Plant Sci.* 11, 1179
5. Vassilev, N. *et al.* (2020) Formulation of microbial inoculants by encapsulation in natural polysaccharides: focus on beneficial properties of carrier additives and derivatives. *Front. Plant Sci.* 11, 1–9 270
6. Sessitsch, A. *et al.* (2019) Microbiome applications from lab to field: facing complexity. *Trends Plant Sci.* 24, 194–198
7. Hou, S. *et al.* (2021) A microbiota–root–shoot circuit favours Arabidopsis growth over defence under suboptimal light. *Nat. Plants* 7, 1078–1092
8. Cosme, M. *et al.* (2018) Non-mycorrhizal plants: the exceptions that prove the rule. *Trends Plant Sci.* 23, 577–587
9. Bulgarelli, D. *et al.* (2015) Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* 17, 392–403
10. Van Deynze, A. *et al.* (2018) Nitrogen fixation in a landrace of maize is supported by a mucilage-associated diazotrophic microbiota. *PLoS Biol.* 16, e2006352
11. Fagorzi, C. *et al.* (2021) Nonadditive transcriptomic signatures of genotype-by-genotype interactions during the initiation of plant–rhizobium symbiosis. *mSystems* 6, e00974-20
12. McCarty, N.S. and Ledesma-Amaro, R. (2019) Synthetic biology tools to engineer microbial communities for biotechnology. *Trends Biotechnol.* 37, 181–197
13. Ludington, W.B. (2022) Higher-order microbiome interactions and how to find them. *Trends Microbiol.* 30, 618–621
14. diCenzo, G.C. *et al.* (2020) Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. *Nat. Commun.* 11, 2574
15. Dini-Andreote, F. and Raaijmakers, J.M. (2018) Embracing community ecology in plant microbiome research. *Trends Plant Sci.* 23, 467–469
16. Busby, P.E. *et al.* (2017) Research priorities for harnessing plant microbiomes in sustainable agriculture. *PLoS Biol.* 15, e2001793

Check for updates

Received: 24 November 2022 | Accepted: 10 January 2023

DOI: 10.1002/sae2.12038

REVIEW ARTICLE

JOURNAL OF SUSTAINABLE AGRICULTURE AND ENVIRONMENT
Open Access
Global Initiative of Crop Microbiome and Sustainable Agriculture

When biodiversity preservation meets biotechnology: The challenge of developing synthetic microbiota for resilient sustainable crop production

Camilla Fagorzi | Iacopo Passeri | Lisa Cangioli | Francesca Vaccaro | Alessio Mengoni

Department of Biology, University of Florence, Firenze, Italy

Correspondence
Camilla Fagorzi and Alessio Mengoni, Department of Biology, University of Florence, via Madonna del piano 6, 50019 Firenze, Italy.
Email: camilla.fagorzi@unifi.it and alessio.mengoni@unifi.it

Funding information
Ministero delle Politiche Agricole Alimentari e Forestali, Grant/Award Number: MICRO4Legumes

Abstract

Agriculture needs to develop novel strategies and practices to meet the increasing global food demand, in an ecological and economical sustainable framework. The plant-associated microbiota is gaining increasing attention as part of these strategies since it strongly contributes to plant health, nutrition, and resilience to environmental perturbations. However, plant domestication has brought to the reduction of the plant abilities to recruit a beneficial microbiota. It is becoming clear that successful use of the plant microbiota requires a multifaceted approach where microbiologist, geneticists, plant scientists, agronomists, and computational biologists can develop ways and solutions to modify both the plant microbiota and plant's ability to recruit it, directed to increase crop performances. Here, while briefly reviewing the state-of-the-art in plant microbiota research, we focus the attention on the need to discover, understand and use the microbiota associated with wild relatives of crops and with neglected crops, which harbour the microbiota biodiversity needed for developing efficient bioinoculant solutions. In particular, we emphasize the convergence of in situ plant biodiversity preservation with microbiome preservation, which provides added value to nature and habitat conservation, as living collections of microbiome biodiversity. The heuristic value of bioinoculants (viz., synthetic communities) and the need of proper computational models to predict the outcome of their applications is also discussed toward a systems-biology-guided synthetic microbiota development.

KEYWORDS

ecosystems services, endophytes, microbiota modeling, plant microbiome, rhizosphere

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2023 The Authors. *Journal of Sustainable Agriculture and Environment* published by Global Initiative of Crop Microbiome and Sustainable Agriculture and John Wiley & Sons Australia, Ltd.

1 | THE NEED OF FEEDING THE WORLD IN A MICROBIOLOGICAL PERSPECTIVE

1.1 | The challenge of sustainable food production

Under current trends, the world population is expected to reach 9.8 billion people by 2050. To meet the increasing demand for food and feed, agricultural production should increase by 50%–69% over the 2010–2050 period. To date, agricultural land area exploitable for food production covers 38% of the global land surface. Two-thirds of this is used as meadows and pastures, while just one-third consists of cropland. From 1961 to the present day, there has been a continuous reduction of global cropland area per capita, from about 0.45 hectares per capita to 0.21 hectares per capita. Increasing global food demand and difficult production conditions, worsened by climate change, is threatening food security situation (Funk & Brown, 2009). Indeed, climate change will impact many elements of the crop production environment; atmospheric carbon dioxide concentration and the average temperature will increase, droughts will be more frequent and severe, intense precipitation events will lead to flooding, and extreme environments will be more common.

The stability of climate experienced during Holocene contributed to the development of agriculture, that is a predictable stable production of food (Richerson et al., 2001). Climate instability, characterized by droughts, floods, cyclones, unseasonal rains, and heat and cold waves, is a major factor limiting crop production. For this reason, we are now facing the challenge to maintain agricultural productivity, trying to reduce climate change effects on human nutrition. Currently, research is aimed at developing new climate-resilient crop varieties and identifying stress-resistant cultivars. Traits characterizing crop adaptation to environmental stresses can be identified through genomic tools, then selected and potentially transferred into specific crops.

Moreover, with the worsening of the climatic conditions, the risk of relying on a few crops, such as for instance rice, wheat, soybean, and maize, could compromise food and nutritional security. One of the strategies should consider widening the portfolio of staple crops, including underutilized crops, with high nutritional content and often adapted to hostile abiotic conditions (Gregory et al., 2019). However, on a complementary side, agrobiotechnologies are forced to find solutions to let present-day crops also more resilient to environmental stress. The crop microbiome, that is the community of microbes inhabiting crop tissues and the plant–soil interface is among the recognized key elements contributing to plant resilience (Trivedi et al., 2020).

Here, while briefly summarizing the state-of-the-art of crop microbiome, we are focusing the attention on the plant microbiome, as a genetic resource, to be preserved and exploited for increasing crop resilience and productivity.

1.2 | The solution of microbial inoculations of plants

The exploitation of the plant and soil microbiome, that is the use of the wide genetic and functional diversity of beneficial microorganisms that colonize plants tissues and the rhizosphere soil, has emerged as an encouraging strategy to face agricultural challenges related to climate change (Qiu et al., 2019). Among these microorganisms, plant growth-promoting microorganisms (PGPM) can be used to formulate consortia able to promote overall plants health and yield (Figure 1). Indeed, plant growth-promoting Rhizobacteria (PGPR), plant growth-promoting bacteria (PGPB), and arbuscular mycorrhizal fungi (AMF) can provide several functions, including biological nitrogen fixation, solubilization of soil phosphorus and potassium, siderophores production, phytohormone (auxin) production and the stimulation of plant defence mechanisms, that mitigate both biotic and abiotic stress.

Traditional bioinoculants are based on a single strain characterized by growth-promoting (PGP) activities. However, while in vitro studies are often giving interesting results, in-field applications produce variable and sometimes unsatisfactory results, indicating the urgency to find solution to limit the lab-to-the field gap (Sessitsch et al., 2019) (Box 1). Suggestions for the need of a systematic approach in developing successful bioinoculant formulation have been indicated (Trivedi et al., 2020). In the last few years, thanks to advancements in multiomics and computational technologies, synthetic microbial communities (SynComs), small consortia of microorganisms designed to promote overall plant productivity as well as outcompete and/or suppress pathogens, have gained momentum (de Souza et al., 2020).

However, despite the progress made, the knowledge gap on detailed mechanistic intra- and inter-kingdom interactions needs to be filled through evaluation of trophic exchanges in the rhizosphere and endosphere (the zone including the internal part) of the plant.

2 | THE PHYTOSPHERE AS A MICROBIAL VAULT OF BIODIVERSITY

2.1 | Genes, taxa and the microbial diversity

Microbes are found everywhere. Moreover, prokaryotic microbes are different from *macrobes* (plants and animals), not only because of obvious biological differences in their cell structure/morphology and physiology, but also because they challenge some basic epistemological notions of natural sciences, as that of species. The biological species concept, *sensu* Mayr, cannot be applied to microbes, since they do not have sexual reproduction, and genomic studies have clarified that a prokaryotic taxon is composed of many different nonoverlapping genes sets (the pangenome concept) (Doolittle & Papke, 2006). This means that when considering an environment, we cannot simply observe a group of taxa, carrying a defined set of genes, but mainly a group of similar genes (i.e., functions) that can be carried by different taxa. Most of these genes are acquired by *h*orizontal gene *t*ransfer (HGT) driven by *m*obile genetic *e*lements (MGEs) when microbes colonize the same environment. We can thus consider the dynamic interplay among hosts, MGEs, and environments as network that shapes functions within communities of an environment (Skippington & Ragan, 2011). Each ecosystem undergoes variation when subjected to stressful factors (i.e., climatic changes) and, typically, the result is a wave of biodiversity loss. The ability to

2767035x, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sae2.12038 by Universita Di Firenze Sistema, Wiley Online Library on [31/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

JOURNAL OF SUSTAINABLE AGRICULTURE AND ENVIRONMENT
Global Initiative of Crop Microbiome and Sustainable Agriculture

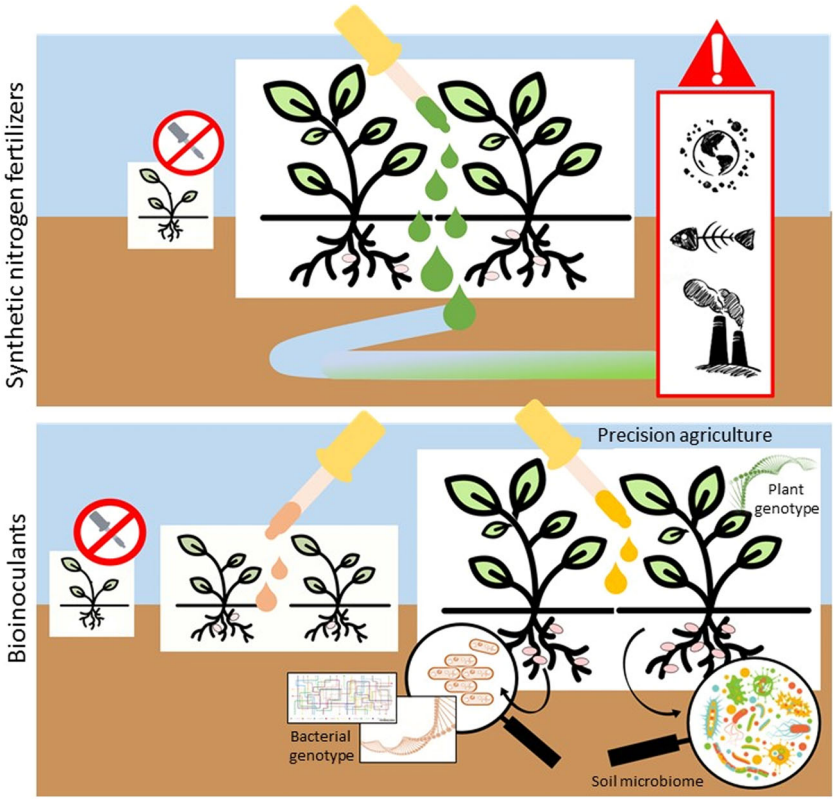


FIGURE 1 Bioinoculants represent a sustainable complement to conventional practices. For instance, bioinoculants with the ability to fix atmospheric nitrogen can complement synthetic nitrogen fertilization and can represent a smart solution in light of precision agriculture, being present and more active in contact with the plant and when the plant needs more nitrogen. Moreover, specificity for plant genotypes and soil types allows to reduce the lab-to-the-field gap, increasing plant colonization and survival in the environment.

maintain or recover its functions in response to external influences, also defined as ecosystem resilience, is connected to the skill of the species to counteract adverse events (Oliver et al., 2015). Functional diversity gives us a mechanistic link between ecosystem and species resilience.

A variegate functional composition of a microbiota should be preserved to make the community capable of facing environmental perturbations (Walker et al., 2022). Some measurables are needed to define the relationship between ecosystem function and diversity. Species richness is questioned, since it implies that all species are potentially equal with respect to function. This concept is clear when talking about predators and plants but should be considered valid also when talking about microorganisms, for the uncoupling between taxa and gene functions (Bengtsson, 1998). The *keystone species* concept, defined as those limited number of organisms and groups of organisms that seem to control the critical processes necessary for ecosystem functioning, as much as it is highly relevant for identification of ecological drivers (Banerjee et al., 2018), does not provide a quantitative measure of the involved processes nor does it describe the link between diversity and function (Banerjee et al., 2018). To relate a measure of diversity to ecosystem function, a possible solution is to analyze the diversity of functional groups (i.e., the number of functional groups) (Moore & de Ruiter, 1991). However, this suffers from some biases, first, the arbitrariness with which functional groups are defined: no identical species exist and species can be observed focusing on different functions. Functional groups can be defined in relation to feeding relationships (e.g., pathotroph, phototroph, osmotroph, and phagotroph), involvement in biogeochemical cycles (e.g., nitrogen-fixers, sulphur oxidizers) but also nontrophic effects, such as nutrient distribution and environmental structure and conditions (e.g., psychrophilic and piezophiles).

Striking examples of the importance of functions over taxa are coming from plant microbiology, for instance, when considering diazotrophs, microbes which carry out atmospheric nitrogen fixation. Symbiotic nitrogen fixers (called rhizobia) are part of the microbiota associated with the root of leguminous plants, where they provide the plant with fixed nitrogen. Rhizobia are a clear example of a functional group of microbes of polyphyletic origin. Rhizobia can be found in both alpha and beta proteobacteria. Additionally, even in the same genus (e.g., *Sinorhizobium*), the functional gene set required for symbiosis and nitrogen fixation (i.e., for being a rhizobium) can be scattered among distantly related strains and also be of different origin (Fagorzi et al., 2020).

As pointed out by M. O'Malley, this uncoupling between taxa and functions has profound consequences on the way we look at life evolution and biodiversity (O'Malley, 2014). In other words, when speaking about biodiversity in microbes, conserving genes can be more relevant than preserving taxa, as it is on the contrary the case for *macrobes* biodiversity, challenging some assumptions we have in biology and nature conservation (O'Malley & Dupré, 2007).

2.2 | Biodiversity in the phytosphere

The microbiota of the phytosphere is known to harbour a wide variety of microbes, including beneficial, neutral, and pathogenic ones. Beneficial microbes are considered a primary resource when

2767035x, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sae2.12038 by Universita Di Firenze Sistema, Wiley Online Library on [31/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

BOX 1. Most pressing questions in plant microbiota research

The high variability in the efficacy of bioinoculants, combined with the lack, in many cases, of agronomic indication on their use, risk producing a lack of confidence in the use of bioinoculants as a complementary or alternative solution to high-input agronomic practices. An effort is therefore needed, both in terms of basic research and technological applications in the field through the transfer of scientific knowledge, to shorten the lab-to-the-field gap (Sessitsch et al., 2019), and to develop successful bioinoculant synthetic microbial communities. This development must be accompanied by providing guidelines to farmers for their use, as well as novel agronomic practices.

In 2017, Busby et al. (2017a). listed five research priorities for plant microbiome research, including novel model systems for understanding the rules of microbiome assembly and the interactions between host plant genotype, microbiome, environment, and management. We think these questions are still valid after 5 years. Progresses have been made, especially on basic knowledge (for a recent review see Trivedi et al., 2020), but, still, the limiting factor is how to incorporate knowledge into management practices. Ideally, a scientific approach similar to drug validation, consisting of a discovery phase and laboratory proofs of efficacy, followed by trials in controlled conditions and in open field, should be applied. All these steps must require consensus on standards (e.g., number of conditions, replicates, yield reporting). We think we need to answer at least five main questions (Figure 2), spanning from basic understanding of the plant microbiomes (Questions 1 and 2), to the interplay among host, microbiome, and environment (Question 3), to the development and use of effective bioinoculants (Questions 4 and 5).

These questions require a complementary work on additional priorities, as (i) the understanding of microbiome assemblage and the prediction of their functionality, (ii) the promotion of culturing plant-associated microbiota, and (iii) the scouting of the microbiome diversity in landraces, nonconventional crops, nonmodel plants. A tight connection with farmers and agronomists is then pivotal to develop novel effective microbiome-based crop management on classical and new crops.

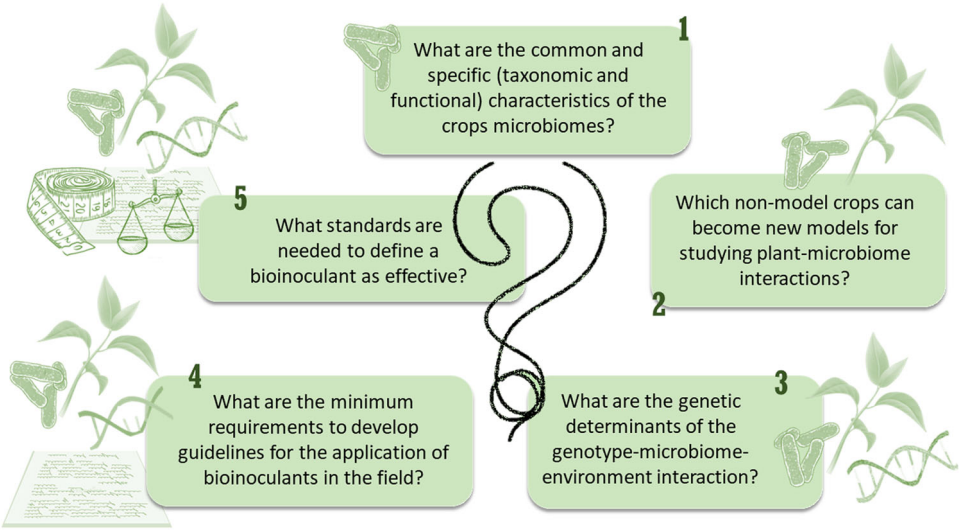


FIGURE 2 Main questions in the understanding and exploiting plant microbiota. Five most pressing questions (Questions 1–5) on research challenges in the plant microbiota and in the agronomic translation of the knowledge are reported.

talking about protection from disease, stress resistance, and growth in biomass. To carry out these functions, microbes act by increasing nutrient mobilization to plants, producing phytohormones, and manipulating plant functional traits, including the immune system (Friesen et al., 2011). Among these microorganisms, bacteria and fungi are predominant. The rhizosphere can contain up to 10^{11} microbial cells per gram of root and more than 30,000 prokaryotic species (Mendes et al., 2011). About a few thousand bacterial and fungal taxa have been reported from plant tissues (Berendsen et al., 2012). The biodiversity of microorganisms in the phytosphere is naturally variable, impacted by continuous environmental variations spanning from roots physiology to the aboveground part growth (e.g., the canopy), strongly affected by aerial dispersion of the microorganisms (Oukala et al., 2021).

Preserving this astounding biodiversity would be ideal but appears to be impossible in reality. What should be preserved? Taxa

2767035x, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sae2.12038 by Universita Di Firenze Sistema, Wiley Online Library on [31/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

or the functions (genetic repertoire) they carry on? Are they all equally relevant? Which rational for the choice of what to preserve should we apply? And, finally, how to preserve since several bacterial taxa are still recalcitrant to cultivation or uncultivable?

Nikolay Vavilov was a central figure in the concept of biodiversity preservation. He created the largest collection of plant germplasm, scouting plants from all over the world. His research led to the identification of the centres of origins of cultivated plants and highlighted the importance of preserving biodiversity. Vavilov's collection is still active at the N.I. Vavilov All Russian Institute of Plant Genetic Resources (VIR) and initiative as the Svalbard Global Seed Vault (https://www.seedvault.no/) are in the frame of Vavilov's legacy.

However, after the introduction of the holobiont concept following the recognition of the large number and great importance of microorganisms associated with plants and animals, it is evident that preservation of seeds is only a part of our safeguard to a future food production. Host plants and their microbes need to be preserved together to guarantee adequate effectiveness of future crop improvement. Concerning the human microbiota initiative for its preservation are ongoing, as "The microbial vault" (https://www.microbiotavault.org/). This initiative is based on biobanking of human faecal samples, microbial isolation and culturing and long-term storage, under the recognition that gut microbiome is the one that mostly impacts human holobiont health.

Similar to human gut microbiota preservation, global efforts toward the analysis and future conservation of microbes associated to the main crops are active (https://www.globalsustainableagriculture.org/). Soil biodiversity hotspots have recently been identified, which should guide prioritization of conservation efforts (Guerra et al., 2022). However, here the challenge is much higher than for human microbiota: we need to preserve the microbiota of many different host species, all of those which can provide food and healthy ecosystems.

Conserving microbes is not an easy task. Many relevant microbes can be nonculturable under currently used conditions and efforts on plant microbiota culturomics initiatives are needed for increasing the culturability and have adequate collections of microorganisms (Mapelli et al., 2022). Though this kind of initiatives can be feasible for some staple crops, they cannot be applied to wild relatives, minor crops, and so on, as the nonmodel plants cited above. A solution is to preserve the plant in situ, on its local environment. Preserving the plant means preserving its microbiome, which is even containing more genetic information than the plant itself. Indeed, we should consider that a single prokaryotic taxon can contain a large genetic diversity, due to a possible open pangenome. Many prokaryotic species colonizing soil and having multiple niches (as soil and plant tissue) normally harbour a large dispensable genome, which determines variable phenotypes with respect to plant interaction (see e.g., Fagorzi et al., 2021). Consequently, though the genetic repertoire needed to interact with plants can be relatively small (Levy et al., 2018), an extensive set of modifier genes, relevant for plant growth promotion can be differentially present in the same taxon. On the other hand, the number of prokaryotic taxa which can be prioritized for such analyses and preservation could be relatively limited. A recent meta-analysis (Riva, 2019) showed that Gammaproteobacteria, Bacilli and Actinobacteria were the predominant groups which constituted core of the plant-associated microbiota, and within them, core genera were *Pseudomonas*, *Pantoea*, *Stenotrophomonas*, *Enterobacter*, *Bacillus*, *Staphylococcus*, *Paenibacillus*, *Microbacterium*, and *Curtobacterium*. However, we cannot exclude that this relatively limited number of taxa simply reflects a limited ability to culture other taxa, prompting to a renewed effort on novel microbial cultural methods (Mapelli et al., 2022).

2.3 | Promoting phytosphere-friendly breeding programs

The process of crop domestication arose around 10,000 years ago, when human civilization became largely dependent on agricultural products, providing a continuous food supply and promoting a sedentary lifestyle (Richerson et al., 2001). Consequently, plant species domestication has significantly contributed to human civilization, but, on the other hand, led to a strong loss of genetic diversity in modern crop cultivars due to a selection of a few species and germplasm, which better adapt and respond to agronomic practices (Evenson & Gollin, 2003). Domestication affected root architecture (i.e., root length and branching), root exudate production and/or composition, and plant defence mechanisms, reducing the ability of the plant to establish beneficial associations with rhizospheric microbes, which may have prominent effects on plant growth and health. Recent studies reported changes in the rhizospheric microbiota of plants such as common bean, lettuce, barley, rice, and *Arabidopsis* with the general trend of diminished Bacteroidetes and increased Proteobacteria and Actinobacteria in comparison to wild relatives (see for instances Bulgarelli et al., 2015; Kim et al., 2020). If we assume that plants partially depend on the microbial community present in the rhizospheric soil as an effect of natural selection, a loss or variation of the traits needed to recruit host-specific root microbiota in modern crop cultivars may have occurred compared to their wild relatives, which are genetically more diverse and adapted to preagricultural soils. For instance, wild plant relatives represent a "living library" of both diverse microbial and plant genetic traits that may have been lost during domestication and can be used as to improve the resistance of modern cultivars to biotic and abiotic stresses (Raaijmakers & Kiers, 2022). Moreover, assessing and accessing the microbiome of indigenous plants and their native habitats represents a new strategy to further exploit microbes and plant traits in modern agriculture. Consequently, molecular breeding and marker-assisted selection can be seen as biotechnological tools to improve beneficial plant–microbe interactions in crop systems and to face an increasing global food demand. This represents one of the major challenges due to the impossibility to expand croplands and the need to reduce the usage of synthetic fertilizers and pesticides.

Since the interaction between host plant and its microbiota is controlled, at least partially, by the plant genome, a prerequisite for the rational manipulation of the plant microbiota for agriculture and

2767035x, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sae2.12038 by Universita Di Firenze Sistema, Wiley Online Library on [31/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

to exploit soil microbes for sustainable crop production is the identification of the plant genes shaping microbiota composition in the rhizosphere. This is particularly relevant for the microbiota associated with wild relatives of the crops, which, having evolved under marginal soil conditions, may represent an untapped resource for low-input agriculture.

Recently, by combining microbiota analysis and quantitative genetics, a locus harbouring a gene encoding an immune receptor with a major effect on the composition of rhizosphere microbiota was identified, confirming the strong link between immunity and beneficial microbiome recruitment (Hacquard et al., 2017) A similar approach was applied to tomato comparing the microbiome of diverse hybrid populations of wild and domesticated cultivars (Oyserman et al., 2022). Here both plant (e.g., iron regulator) and bacterial traits (polysaccharides, vitamins, sulphur, and iron metabolisms) associated to tomato quantitative trait loci (QTLs) were found. A link between the genetics plant and its microbiome have also been in the understanding of the heterosis in maize, a still not-well clear phenomenon called also hybrid vigor, where first-generation offspring outperform their parents (Wagner et al., 2021). These studies demonstrated the power of an integrated approach based on bacterial and plant genomics and directions for plant breeding programs and selection of microbial functions. QTLs can be targeted to obtain improved plant–microbe interactions, identifying "missing microbes" or functions lost during the domestication process, and provide model studies to dissect the molecular details driving plant–bacteria interactions.

2.4 | Wild relatives and nonmodel plants: Microbial biodiversity hotspots?

One of the greatest challenges to face when putting into practices the knowledge of the plant microbiomes is defining methods and principles to translate the results acquired on model plant species to real crops. Often real crops are nonmodel species or genotypes different from model systems (Busby et al., 2017b). Ideally, to allow a smoother translation of knowledge and methods, agricultural germplasms for which close relatives exist as model species, should be prioritized (Vaccaro et al., 2022). Crops, such as alfalfa, chickpea, lentil, landraces of rice, maize, and wheat can represent an excellent opportunity to take advantage of, since they hold their own specific features, but at the same time they can have relatives as model species (such as *Medicago truncatula* for alfalfa) (Vaccaro et al., 2022). Since wild relatives of crops evolved under marginal soil conditions, the microbiome associated with them is expected to harbour functions helping the plant to cope with low nutrients inputs and environmental heterogeneity, two conditions normally found in natural habitats. Because of that their microbiome can represent an untapped resource for the challenges presented by low-input agriculture (Escudero-Martinez & Bulgarelli, 2019). Moreover, microbiomes of wild relatives and nonmodel species, represent the benchwork to investigate the structure and the evolution of the plant-associated microbial communities. Then, they may provide mechanistic insights which could be ultimately useful and critical to rationally design synthetic microbial communities (Vaccaro et al., 2022).

Examples of insights gained from nonmodel plant species can be found in the identification of novel genetic determinants of host–bacteria interactions. A study on local crop varieties of alfalfa (*Medicago sativa*) revealed the existence of genotypic interactions influencing the phenotypic traits of the interaction with the rhizobial nitrogen-fixing symbiont *Sinorhizobium meliloti* (Fagorzi et al., 2021). Here, many genes (up to more than 1000 on a single strain) were influenced in their expression by the strain genotype (35%), by the genotype of the variety of host plant tested (16%), and by strain-by-host plant genotype interactions (29%). These data indicate from one side that alfalfa can largely benefit from breeding programs considering the selectivity of the symbiotic rhizobia recruitment, and, on the other side, that local varieties can host a large genetic diversity in their symbiotic microbiome. Another study on a Mexican indigenous maize landrace shed light on the vault of microbial biodiversity hosted in nonmodel plants. Here, it was shown that the microbiota associated to the mucilage secreted by aerial roots provides unexpected biological nitrogen fixation abilities which contribute to plant nutrition (Van Deynze et al., 2018). Since diazotrophic microbial associations with cereals is stirring a large interest in low-input agriculture, this example highlights the importance of research on nonmodel plants for discovering novel microbial functions and planning novel microbe-oriented plant.

In our opinion, research on landraces and consequently their preservation can be highly important not only to save important germplasm for breeding, *sensu* Vavilov, but also to save the plant-associated microbial diversity (which is yet to be investigated in such plants) and which can harbour functions for sustainable food production. For instance, concerning climate-resilient nonmodel crops of high nutritional importance, such as *Eragrostis tef* (teff), *Chenopodium quinoa* (quinoa), *Setaria italica* (foxtail millet), *Amaranthus* spp., no publication on their native microbiota is available which should prompt prioritization in future research programs (Gregory et al., 2019).

In conclusion, we should shift from plant preservation to holobiont preservation and here is when biodiversity and habitat preservation meets biotechnology (as exploitation of genetic resources). In our opinion, this concept of habitat conservation which includes the microbiome, must be considered when biodiversity reserves are planned and managed as well as when traditional, local or neglected crops are valorized in programs at the FAO Globally Important Agricultural Heritage Systems (GIAHS) (https://www.fao.org/giahs/). However, for a rational decision of (nonmodel) plant holobiont biodiversity hotspots studies on their microbiome must increase, as to date information is scarce.

3 | BIOTECHNOLOGY ON BIODIVERSITY: TOWARD THE UNDERSTANDING OF THE RULES OF ENGAGEMENT OF SYNTHETIC MICROBIOTA

As previously discussed, the relationship between microorganisms and plants is not a binary host-microbial interaction but a complex community assembly, summarized by the concept of holobiont, which

2767035x, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sae2.12038 by Universita Di Firenze Sistema, Wiley Online Library on [31/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

considers the emerging interactions and outcomes of hosts and their microbiota (Theis et al., 2016). In response to such complexity, the process of bioinoculants design has been driven by new approaches. *Syn*thetic microbial *Com*munities (SynComs) relies on microbial ecology concepts, but their specific design balances between well-established theories of ecology and the need of predictive modelling to face the paucity of specific physiological data on microbes' functions and host-microbes cross talks. The general idea is to take advantage from the presence of a core plant microbiota and create a consortium of (several) strains which can mimic the taxonomy, functions, and effects of a healthy native plant microbiota. Several formulations have been tested to enhance yield and nutrient content of different crops, confirming the importance of the development of SynComs for agricultural use (Maitra et al., 2022). However, various strategies need to be applied to effectively engineer plant microbiomes, working either on plant side or on the microbial side (French et al., 2021). Moreover, abiotic factors (e.g., physicochemical soil characteristics) must be considered, since they can act directly on the microbial side by determining, for instance, the persistence in soil and microbial activity, and indirectly, by changing amount and type of plant-secreted chemical compounds. Recent computational methods, including mathematical models, Machine Learning and AI, are strengthening the development of efficient inoculants for plant's benefits (Leggieri et al., 2021; de Souza et al., 2020; Vorholt et al., 2017). These methods help to make use of the still limited knowledge about the interactions between microbes and plants and to take into consideration the many variables of the system (either biotic or abiotic). The most common methods which can be used to rationally design SynCom, span from statistical methods based on taxa occurrences to methods relying on metabolic exchanges (Figure 3).

Models of inter-species interaction from theoretical ecology, as Resource Ratio Theory (RRT), Maximum Power Principle (MPP), and Black Queen Theory (BQT), have been successfully employed to analyze interspecies microbial interactions (Dunham, 2007). The RRT (Figure 3a) is a model that could explain the conditions under which natural selection may favour cooperation over competition. Moreover, theories dealing with mutualism have spread out in recent years, including game theory (see next paragraph). MPP (Figure 3b) may provide predictions on communities clustering benefits looking for maximum power outcomes in ecological situations, while BQT can explain genome evolution in symbiotic bacteria (Morris et al., 2012).

Game theory (GT) is another general mathematical framework to model strategic interactions among a number of agents (players) where the payoff of each agent is not only a function of its own strategy (action) but also a function of other players' strategies. Cooperation and competition have a fitness face-off in game theory. Interactions among players are costly, the energy spent is a cost (loss) and the resources are benefits (gain): the net between gain and loss is the payoff. In GT the payoff for any player depends not only on its own strategy, but the strategies of the other players in the environment (Figure 3c). GT can help to understand the evolution of mutualism between microbes and hosts and, by establishing a cost for the mutualism, how much we can gain increasing the microbial counterpart (Leggieri et al., 2021; Zomorrodi & Segrè, 2017).

Still concerning mathematical frameworks, population dynamic models, as the generalized Lotka-Volterra (LV) models have been used to model bacteria-bacteriophage interactions, microbial gut interactions, and microbial communities' interactions, as models to predict outcomes in plant–microbe interactions (Vorholt et al., 2017).

However, we should consider that natural microbial communities are characterized by a high-level complex in their structure. The self-organization of the community itself depends on factors such as the heterogeneity of the habitat, natural gradients, interaction abundances, and many others. Spatial models, where the dynamics of the system across different locations is studied with partial differential equations, can be applied in such cases. The most widely used model is the reaction-diffusion equation, which determines the density of each species at different times and different locations in space caused by diffusion and population dynamics. This class of models has been used in combination to population-based models to study a variety of ecological phenomena related to spatial effects such as range expansion and diffusion-based spatial patterning. Recently, spatial models have been used to understand and predict the root-to-shoot circuit in *Arabidopsis* (Hou et al., 2021), where, under low light, leaves induce a modulation of the rhizosphere microbiota through physiological changes in roots. Altered microbiota, in turn, responds by alleviating plant stress due to low photosynthetic activity.

Going to prediction of taxa and genes (i.e., metabolic functions) relevance to the assembly and functionality of the microbiome, generalized linear models, as the Poisson regression, are well-established methods. The Poisson regression model is a technique used to describe count data as a function of a set of predictor variables. The Poisson distribution occupies a central position in both ecological theory and practice and allows to model time-series data in microbial ecology. Poisson regression has been used to predict plant status from rhizosphere microbiota (Bacci et al., 2018), providing the basis for the identification of relevant "predictor" taxa of the outcomes which can be prioritized in SynCom constitution (Hou et al., 2021).

Machine learning (ML) is another relevant technology for making sense of multi-omics data on plant-associated microbiome and to detect insightful data patterns (Figure 3d). ML can be used to understand how different organisms in a microbial community affect each other's growth to form complex networks of interdependencies. For example, Random Forest (RF) can be used to predict microbial interactions given trait descriptions (features) for each member of the community or extrapolate trait combinations that are causally relevant to species interactions (DiMucci et al., 2018). ML has been variously used in host–microbiota interaction, for instance, to link host plant (rice) functional traits to the core microbial community (Guo et al., 2021).

Finally, coming to the possibility to predict the metabolic outcomes from genome or a metagenome, metabolic *gen*ome-scale network *re*constructions (GENREs) are becoming popular. GENREs are intracellular network mathematical models consisting of a compilation of all biochemical reactions occurring in an organism (Fang et al., 2020). These models start from genome annotation and defining a stoichiometric matrix, are forced to solution by constraints (e.g., mass balance, steady-state conditions, etc.), and typically by applying a biomass-producing

2767035x, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sae2.12038 by Universita Di Firenze Sistema, Wiley Online Library on [31/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

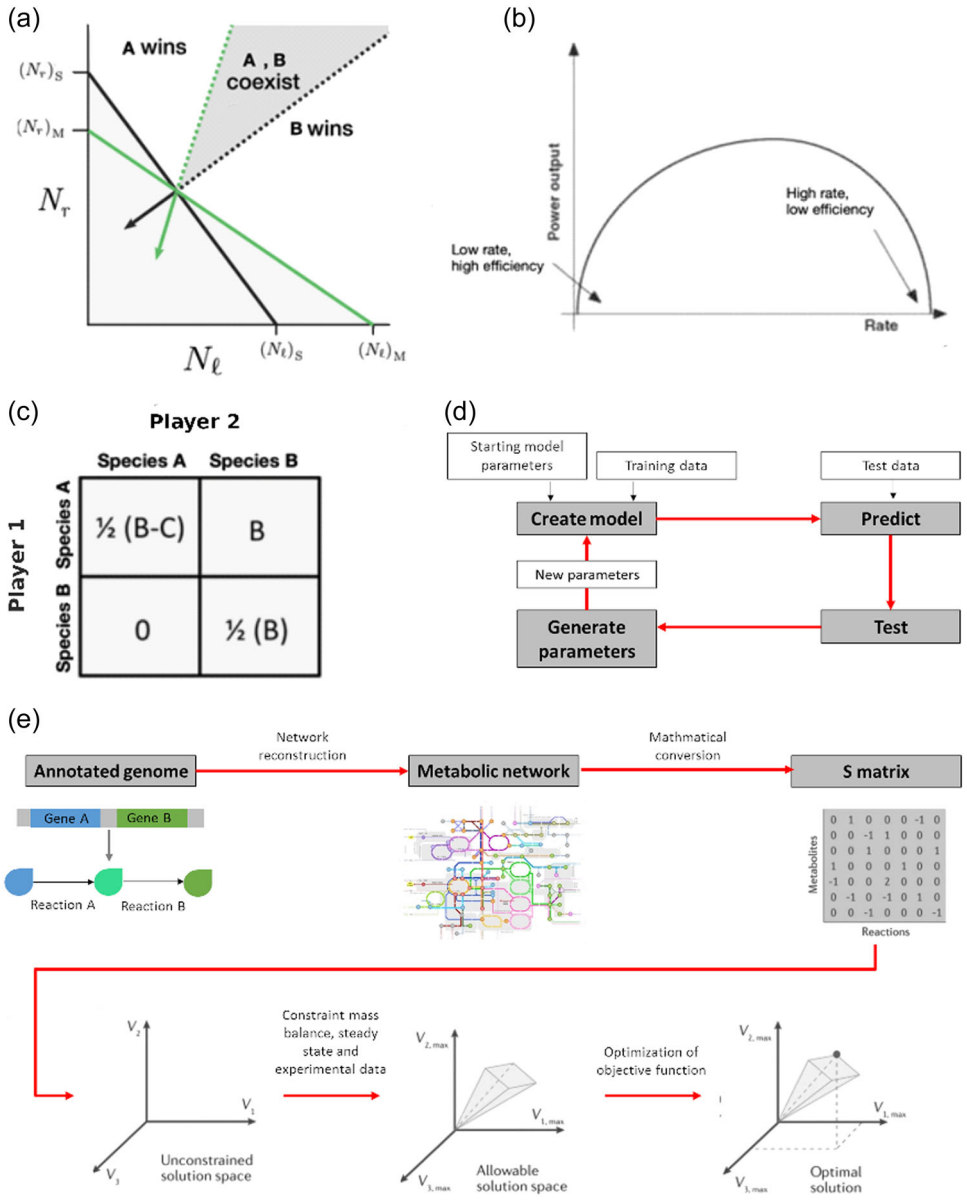


FIGURE 3 Developing model to investigate population dynamics and metabolic networks. (a) Scheme for the RRT modelling the competition between two or more species (A and B) for a limiting resource; prediction of outcomes including coexistence of species can be derived from single species abundance and rates of resource utilization; application of this modelling spans from conflicts in symbiosis (e.g., symbiotic nitrogen-fixing microbiota) to social interaction among members of the microbiota. (b) MPP is a thermodynamic-based model, where systems self-organize to increase energetic quantities (i.e., metabolic rate); optimum of metabolic fluxes between microbiota and the plant can be predicted in this way; (c) GT models strategic interactions among players where the payoff depends on the strategies of both players. Payoffs for two species: B "benefits," and C "costs," allowing to predict the cost of sustaining a microbiota with respect to the gain in plant biomass. (d) ML workflow starting from a training data set (e.g., in vitro plant-microbe interaction studies) used to create a first model, then tested on real data, which turns out in improved model parameters used to update the model in a recursive way; applications of ML includes studies of association between ecosystem functions and members of a microbiome. (e) Example of workflow in GENREs, where genome annotation provides the basis for the metabolic network reconstruction and the following stoichiometric (S) matrix, which is then solved applying constrains and an objective function to estimate the overall fluxes of reactions; modified from (Fang et al., 2020).

reaction (Figure 3e). The prediction of fluxes for each reaction in the model is an indication of relative importance of reactions and pathways to cell's growth capacity. Application of this modelling to communities is still challenging; however, predictions of metabolic exchanges in simplified systems, such as the rhizobium-legume symbiosis (diCenzo et al., 2020) have been done. For the rhizobium-legume symbiosis, the metabolic model reconstruction of the exchanges in the root nodule (the nitrogen-fixing symbiotic structure) allowed to predict the trade-off values of fixed nitrogen, plant biomass, and root nodule biomass to estimate the maximum nitrogen-fixing capacity of the holobiont.

2767035x, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sae2.12038 by Universita Di Firenze Sistema, Wiley Online Library on [31/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

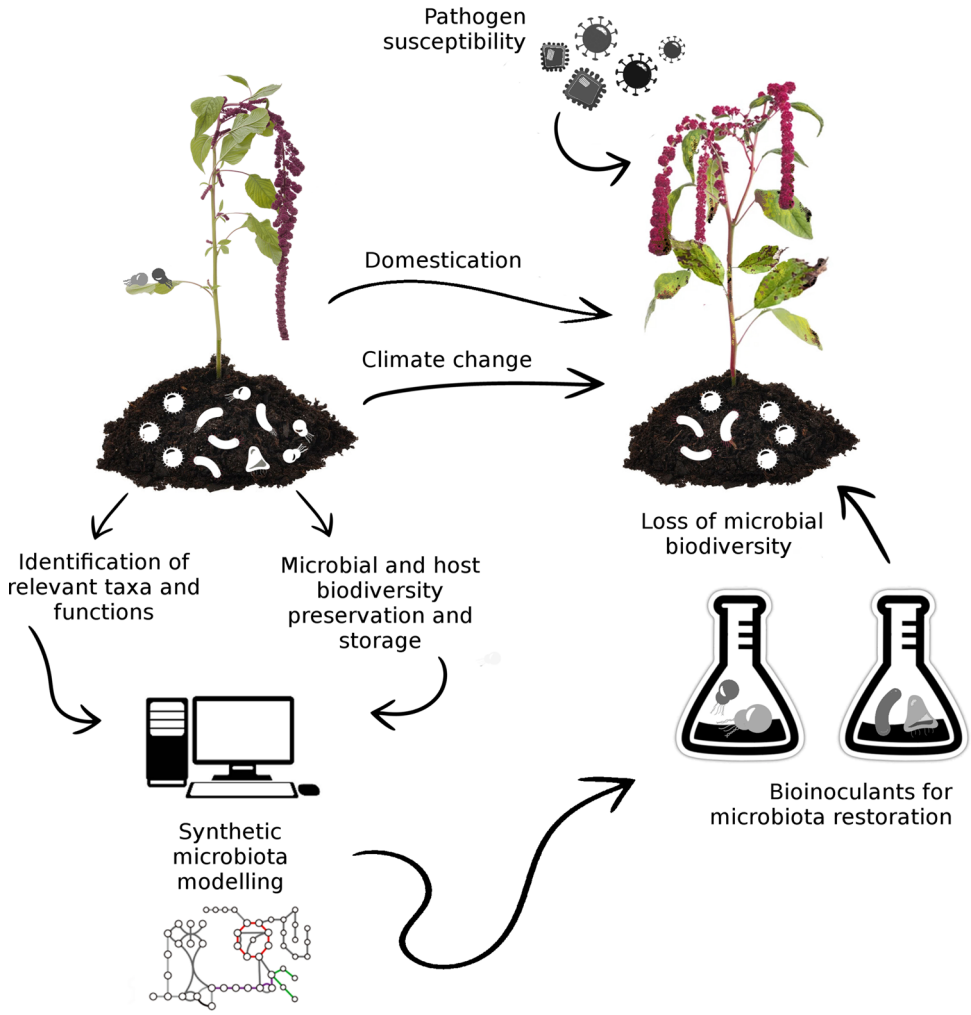


FIGURE 4 Bioinoculants development in a system biology approach. Natural environments represent a rich reservoir of genes and related functions that can be lost during domestication processes and as an effect of climate changes. The study, preservation, and storage of the microbial and host biodiversity lay the groundwork for the development of bioinoculants for microbiota restoration in case of augmented pathogen susceptibility or stress-related effect on the plant. Data obtained from microbiota and functions present in the natural soil can be integrated to create metabolic models that allow a synthetic biology approach to test hypotheses on plant-microbe interaction.

Although the number of approaches, briefly reviewed above, is allowing to dissect interactions among different (micro) organisms and the environment, to rationally design synthetic communities, these methods are not widely used. From one side, the expertize needed to master such models is not commonly found among microbiologists, re-proposing the classical division between lab scientists and computational scientists. Such division can be overcome by introducing more computational skills in educational programs in biological disciplines and by the recognition of the need of interdisciplinarity in research teams. On the other side, we should recognize that, despite a high number of reports and experiments performed on plant-associated microbiomes, relatively few are based on experiments designed to produce data which can be reused in clean computational models. The lack of shared standards, but also the enormous number of variables inherent to biological investigations (i.e., the many environmental variables) are serious limitations. The need to have a consensus on how data are reported is urgent and proposals have been presented (Pacheco et al., 2022), while we still need more data on real crops (often from nonmodel plant species/genotypes) in field and controlled conditions (Vaccaro et al., 2022).

4 | CONCLUSIONS AND FUTURE DIRECTIONS

The intense agricultural production, combined with land degradation, as a result also of global warming and climate alterations, is leading to a general depauperating of microbial genetic resources present in agroecosystems. This decrease in microbial diversity is having serious consequences, including reduction of soil fertility and increased sensitivity of plants to biotic and abiotic stresses. Soil health is a key factor for high crop productivity and requires conscious management of organic and inorganic elements. The integration of natural components such as bioinoculants improves efficiency of agronomic management (Maitra et al., 2022; de Souza et al., 2020). Indeed, bioinoculants

2767035x, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sae2.12038 by Universita Di Firenze Sistema, Wiley Online Library on [31/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

are a core component among the Green Revolution Technologies because of the positive evidence shown by the plants teamed with microorganisms, which enhance the assimilation of nutrients such as nitrogen, resulting in an amelioration of plant's health (Maitra et al., 2022). However, the design of bioinoculants is still facing difficulties, due to our lack of knowledge of the microbes inhabiting the phytosphere, including their taxonomy, functions, and interactions. Programs for bioinoculant development should rely on various data and competences, from-omics data, to culturing and computational biology predictions. Such programs should also take into consideration the holobiont view of a plant, where plant genotypes and microbiome genotypes are interacting to provide increased plant functional traits and consequently that plant microbiomes are plant-genotype dependent (Figure 4). In the perspective of preserving microbes for future applications, we must strongly consider that wild relatives of crops, and wild plants, often harbour the most important part of the microbiomes needed for increased plant resilience to environmental perturbations. For sustainable agriculture, we need then to promote a novel approach to conservation biology, as the conservation of the holobiont. Conserving the holobiont means viewing the conservation of an organism (the plant in our case) as the conservation of its associated microorganisms also, which can be key for the agriculture of the future. In turn, conserving the holobiont means preserving the environment where the holobiont lives, including microbial endemism, since plant-associated microbes change under ex situ growth conditions.

AUTHOR CONTRIBUTIONS

Camilla Fagorzi and Alessio Mengoni conceived the idea. Camilla Fagorzi was responsible for organizing the whole manuscript, contributing to drafting Section 2 and finalizing figures. Lisa Cangioli contributed to drafting Section 2. Francesca Vaccaro contributed to drafting Section 1. Iacopo Passeri contributed to drafting Section 3. Camilla Fagorzi and Alessio Mengoni collected contributions from co-authors and finalized the writing of the manuscript. Alessio Mengoni secured funding and provided guidance to undertake the work. All authors provided feedback on the manuscript.

ACKNOWLEDGEMENTS

This study was funded by MICRO4Legumes, D.M.n. 89267 (Italian Ministry of Agriculture) to A.M. I.P., L.C., and F.V. are supported by a PhD fellowship from D.M.n. 351/2022 (Italian Ministry of University and Research), MICRO4Legumes (Italian Ministry of Agriculture), and PON Ricerca e Innovazione 2014–2020 (Italian Ministry of University and Research), respectively. C.F. is supported by a post-doctoral fellowship from the LEGU-MED project (PRIMA Foundation) Call 2019.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

No data are reported.

ETHICS STATEMENT

The authors declare that they followed the ethics policies of the journal.

ORCID

Alessio Mengoni http://orcid.org/0000-0002-1265-8251

REFERENCES

Bacci G, Cerri M, Lastrucci L, Ferranti F, Ferri V, Foggi B, et al. Applying predictive models to decipher rhizobacterial modifications in common reed die-back affected populations. Sci Total Environ. 2018;642:708–22.

Banerjee S, Schlaeppi K, van der Heijden MGA. Keystone taxa as drivers of microbiome structure and functioning. Nat Rev Microbiol. 2018;16(9):567–76. Available from: https://doi.org/10.1038/s41579-018-0024-1

Bengtsson J. Which species? What kind of diversity? Which ecosystem function? Some problems in studies of relations between biodiversity and ecosystem function. Appl Soil Ecol. 1998;10:191–9.

Berendsen RL, Pieterse CMJ, Bakker PAHM. The rhizosphere microbiome and plant health. Trends Plant Sci. 2012;17(8):478–86.

Bulgarelli D, Garrido-Oter R, Münch PC, Weiman A, Dröge J, Pan Y, et al. Structure and function of the bacterial root microbiota in wild and domesticated barley. Cell Host Microbe. 2015;17(3):392–403.

Busby PE, Soman C, Wagner MR, Friesen ML, Kremer J, Bennett A, et al. Research priorities for harnessing plant microbiomes in sustainable agriculture. PLoS Biol. 2017a;15(3):e2001793.

Busby PE, Soman C, Wagner MR, Friesen ML, Kremer J, Bennett A, et al. Research priorities for harnessing plant microbiomes in sustainable agriculture. PLoS Biol. 2017b;15(3):e2001793.

diCenzo GC, Tesi M, Pfau T, Mengoni A, Fondi M. Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. Nat Commun. 2020;11:2574.

DiMucci D, Kon M, Segrè D. Machine learning reveals missing edges and putative interaction mechanisms in microbial ecosystem networks. mSystems. 2018;3:e00181–18.

Doolittle WF, Papke RT. Genomics and the bacterial species problem. Genome Biol. 2006;7:116.

Dunham MJ. Synthetic ecology: a model system for cooperation. Proceedings of the National Academy of Sciences. 2007;104:1741–2.

Escudero-Martinez C, Bulgarelli D. Tracing the evolutionary routes of plant–microbiota interactions. Curr Opin Microbiol. 2019;49:34–40.

Evenson RE, Gollin D. Assessing the impact of the Green Revolution, 1960 to 2000. Science. 2003;300(5620):758–62.

Fagorzi C, Ilie A, Decorosi F, Cangioli L, Viti C, Mengoni A, et al. Symbiotic and nonsymbiotic members of the genus ensifer (syn. Sinorhizobium) are separated into two clades based on comparative genomics and high-throughput phenotyping. Genome Biol Evol. 2020;12:2521–34.

Fagorzi C, Bacci G, Huang R, Cangioli L, Checcucci A, Fini M, et al. Nonadditive transcriptomic signatures of genotype-by-genotype interactions during the initiation of plant-rhizobium symbiosis. mSystems. 2021;6(1):1–15.

Fang X, Lloyd CJ, Palsson BO. Reconstructing organisms in silico: genome-scale models and their emerging applications. Nat Rev Microbiol. 2020;18(12):731–43. Available from: https://doi.org/10.1038/s41579-020-00440-4

French E, Kaplan I, Iyer-Pascuzzi A, Nakatsu CH, Enders L. Emerging strategies for precision microbiome management in diverse agroecosystems. Nature Plants. 2021;7:256–67.

Friesen ML, Porter SS, Stark SC, Von Wettberg EJ, Sachs JL, Martinez-Romero E. Microbially mediated plant functional traits. Annual Rev Ecol Evol Systematics. 2011;42:23–46.

2767035x, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sae2.12038 by Universita Di Firenze Sistema, Wiley Online Library on [31/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Funk CC, Brown ME. Declining global per capita agricultural production and warming oceans threaten food security. Food Security. 2009;1:271–89.
Gregory PJ, Mayes S, Hui CH, Jahanshiri E, Julkifle A, Kuppusamy G, et al. Crops for the future (CFF): an overview of research efforts in the adoption of underutilised species. Planta. 2019;250(3):979–88.
Guerra CA, Berdugo M, Eldridge DJ, Eisenhauer N, Singh BK, Cui H, et al. Global hotspots for soil nature conservation. Nature. 2022;610(7933): 693–8.
Guo J, Ling N, Li Y, Li K, Ning H, Shen Q, et al. Seed-borne, endospheric and rhizospheric core microbiota as predictors of plant functional traits across rice cultivars are dominated by deterministic processes. New Phytol. 2021;230:2047–60.
Hacquard S, Spaepen S, Garrido-Oter R, Schulze-Lefert P. Interplay between innate immunity and the plant microbiota. Annu Rev Phytopathol. 2017;55:565–89.
Hou S, Thiergart T, Vannier N, Mesny F, Ziegler J, Pickel B, et al. A microbiota–root–shoot circuit favours Arabidopsis growth over defence under suboptimal light. Nature Plants. 2021;7:1078–92.
Kim H, Lee KK, Jeon J, Harris WA, Lee YH. Domestication of Oryza species eco-evolutionarily shapes bacterial and fungal communities in rice seed. Microbiome. 2020;8:20.
Leggieri PA, Liu Y, Hayes M, Connors B, Seppälä S, O'Malley MA, et al. Integrating systems and synthetic biology to understand and engineer microbiomes. Annu Rev Biomed Eng. 2021;23:169–201.
Levy A, Salas Gonzalez I, Mittelviefhaus M, Clingenpeel S, Herrera Paredes S, Miao J, et al. Genomic features of bacterial adaptation to plants. Nature Genet. 2018;50:138–50.
Maitra S, Brestic M, Bhadra P, Shankar T, Praharaj S, Palai JB, et al. Bioinoculants—natural biological resources for sustainable plant production. Microorganisms. 2022;10(1):51.
Mapelli F, Mengoni A, Riva V, Borin S. Bacterial culturing is crucial to boost sustainable agriculture. Trends Microbiol. 2022;31(1):1–4.
Mendes R, Kruijt M, De Bruijn I, Dekkers E, Van Der Voort M, Schneider JHM, et al. Deciphering the rhizosphere microbiome for disease-suppressive bacteria. Science. 2011;332:1097–100.
Moore JC, de Ruiter PC. Temporal and spatial heterogeneity of trophic interactions within below-ground food webs. Agric Ecosyst Environ. 1991;34:371–97.
Morris JJ, Lenski RE, Zinser ER. The black queen hypothesis: evolution of dependencies through adaptive gene loss. mBio. 2012;3:e00036–12.
Oliver TH, Heard MS, Isaac NJB, Roy DB, Procter D, Eigenbrod F, et al. Biodiversity and resilience of ecosystem functions. Trends Ecol Evolut. 2015;30:673–84.
O'Malley M. Philosophy of microbiology. Cambridge University Press; 2014.
O'Malley MA, Dupré J. Size doesn't matter: towards a more inclusive philosophy of biology. Biol Philos. 2007;22:155–91.
Oukala N, Aissat K, Pastor V. Bacterial endophytes: the hidden actor in plant immune responses against biotic stress. Plants. 2021;10:1012.
Oyserman BO, Flores SS, Griffioen T, Pan X, van der Wijk E, Pronk L, et al. Disentangling the genetic basis of rhizosphere microbiome assembly in tomato. Nat Commun. 2022;13(1):3228.
Pacheco AR, Pauvert C, Kishore D, Segrè D. Toward FAIR representations of microbial interactions. mSystems. 2022;7(5). https://journals.asm.org/doi/10.1128/msystems.00659-22
Qiu Z, Egidi E, Liu H, Kaur S, Singh BK. New frontiers in agriculture productivity: optimised microbial inoculants and in situ microbiome engineering. Biotech Adv. 2019;37(6):107371.
Raaijmakers JM, Kiers ET. Rewilding plant microbiomes. Science. 2022;378(6620):599–600.
Richerson PJ, Boyd R, Bettinger RL. Was agriculture impossible during the pleistocene but mandatory during the holocene? A climate change hypothesis. Am Antiq. 2001;66:387–411.
Riva V. A Meta-Analysis approach to defining the culturable core of plant endophytic bacterial communities. Appl Env Microbiol. 2019;88(13): 1–10.
Sessitsch A, Pfaffenbichler N, Mitter B. Microbiome applications from lab to field: facing complexity. Trends Plant Sci. 2019;24(3): 194–8.
Skippington E, Ragan MA. Lateral genetic transfer and the construction of genetic exchange communities. FEMS Microbiol Rev. 2011;35: 707–35.
de Souza RSC, Armanhi JSL, Arruda P. From microbiome to traits: designing synthetic microbial communities for improved crop resiliency. Front Plant Sci. 2020;11(August):1–7.
Theis KR, Dheilly NM, Klassen JL, Brucker RM, Baines JF, Bosch TCG, et al. Getting the hologenome concept right: an eco-evolutionary framework for hosts and their microbiomes. mSystems. 2016;1: e00028–16.
Trivedi P, Leach JE, Tringe SG, Sa T, Singh BK. Plant–microbiome interactions: from community assembly to plant health. Nat Rev Microbiol. 2020;18:607–21.
Vaccaro F, Cangioli L, Mengoni A, Fagorzi C. Synthetic plant microbiota challenges in nonmodel species. TIM. 2022;30(10):922–4. Available from: https://doi.org/10.1016/j.tim.2022.06.006
Van Deynze A, Zamora P, Delaux P-M, Heitmann C, Jayaraman D, Rajasekar S, et al. Nitrogen fixation in a landrace of maize is supported by a mucilage-associated diazotrophic microbiota. PLoS Biol. 2018;16(8):e2006352. Available from: http://dx.plos.org/10.1371/journal.pbio.2006352
Vorholt JA, Vogel C, Carlström CI, Müller DB. Establishing causality: opportunities of synthetic communities for plant microbiome research. Cell Host Microbe. 2017;22(2):142–55.
Wagner MR, Tang C, Salvato F, Clouse KM, Bartlett A, Vintila S, et al. Microbe-dependent heterosis in maize. Proc Natl Acad Sci USA. 2021;118(30):1–8.
Walker JR, Woods AC, Pierce MK, Steichen JL, Quigg A, Kaiser K, et al. Functionally diverse microbial communities show resilience in response to a record-breaking rain event. ISME Commun. 2022;2(1):81.
Zomorrodi AR, Segrè D. Genome-driven evolutionary game theory helps understand the rise of metabolic interdependencies in microbial communities. Nat Commun. 2017;8:1563.

How to cite this article: Fagorzi C, Passeri I, Cangioli L, Vaccaro F, Mengoni A. When biodiversity preservation meets biotechnology: the challenge of developing synthetic microbiota for resilient sustainable crop production. J Sustain Agric Environ. 2023;2:5–15. https://doi.org/10.1002/sae2.12038

2767035x, 2023, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sae2.12038 by Universita Di Firenze Sistema, Wiley Online Library on [31/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

microorganisms

MDPI

Article

Endophytes from African Rice (*Oryza glaberrima* L.) Efficiently Colonize Asian Rice (*Oryza sativa* L.) Stimulating the Activity of Its Antioxidant Enzymes and Increasing the Content of Nitrogen, Carbon, and Chlorophyll

Carmen Bianco [1,*], Anna Andreozzi [1], Silvia Romano [1], Camilla Fagorzi [2], Lisa Cangioli [2], Pilar Prieto [3], Fousseyni Cisse [4], Oumar Niangado [5], Amadou Sidibé [4], Silvia Pianezze [6,7], Matteo Perini [6], Alessio Mengoni [2] and Roberto Defez [1]

1 Institute of Biosciences and BioResources, Via P. Castellino 111, 80131 Naples, Italy; anna.andreozzi@ibbr.cnr.it (A.A.); romanosilvia067@gmail.com (S.R.); roberto.defez@ibbr.cnr.it (R.D.)
2 Department of Biology, University of Florence, Via Madonna del Piano 6, 50019 Sesto Fiorentino, Italy; camilla.fagorzi@unifi.it (C.F.); lisa.cangioli@unifi.it (L.C.); alessio.mengoni@unifi.it (A.M.)
3 Departamento de Mejora Genética, Campus 'Alamedadel Obispo', Instituto de Agricultura Sostenible (IAS), Consejo Superior de Investigaciones Científicas (CSIC), Avd. Menéndez Pidal s/n, 14004 Córdoba, Spain; pilar.prieto@ias.csic.es
4 Institut d'Economie Rurale, Rue Mohamed V Bamako, Bamako B.P. 258, Mali; fousscisse@yahoo.fr (F.C.); amadousidibe57@yahoo.fr (A.S.)
5 Syngenta Foundation for Sustainable Agriculture, Bamako B.P.E. 1449, Mali; oumar.niangado@syngenta.com
6 Fondazione Edmund Mach, Via Mach 1, 38098 San Michele All'Adige, Italy; silvia.pianezze@fmach.it (S.P.); matteo.perini@fmach.it (M.P.)
7 Environmental and Animal Sciences DI4A, Università degli Studi di Udine, Via Sondrio 2/A, 33100 Udine, Italy
* Correspondence: carmen.bianco@ibbr.cnr.it; Tel.: +39-081-613-2610

Abstract: Bacterial endophytes support the adaptation of host plants to harsh environments. In this study, culturable bacterial endophytes were isolated from the African rice *Oryza glaberrima* L., which is well-adapted to grow with poor external inputs in the tropical region of Mali. Among these, six N-fixer strains were used to inoculate *O. glaberrima* RAM133 and the Asian rice *O. sativa* L. cv. Baldo, selected for growth in temperate climates. The colonization efficiency and the N-fixing activity were evaluated and compared for the two rice varieties. *Oryza sativa*-inoculated plants showed a fairly good colonization efficiency and nitrogenase activity. The inoculation of *Oryza sativa* with the strains *Klebsiella pasteurii* BDA134-6 and *Phytobacter diazotrophicus* BDA59-3 led to the highest nitrogenase activity. In addition, the inoculation of 'Baldo' plants with the strain *P. diazotrophicus* BDA59-3 led to a significant increase in nitrogen, carbon and chlorophyll content. Finally, 'Baldo' plants inoculated with *Kl. pasteurii* BDA134-6 showed the induction of antioxidant enzymes activity and the maintenance of nitrogen-fixation under salt stress as compared to the unstressed controls. As these endophytes efficiently colonize high-yielding crop varieties grown in cold temperate climates, they become good candidates to promote their growth under unfavorable conditions.

Keywords: nitrogen fixation; indole-3-acetic acid (IAA); 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase; rice endophytic bacteria; salt stress; anti-oxidative enzymes

check for updates

Citation: Bianco, C.; Andreozzi, A.; Romano, S.; Fagorzi, C.; Cangioli, L.; Prieto, P.; Cisse, F.; Niangado, O.; Sidibé, A.; Pianezze, S.; et al. Endophytes from African Rice (*Oryza glaberrima* L.) Efficiently Colonize Asian Rice (*Oryza sativa* L.) Stimulating the Activity of Its Antioxidant Enzymes and Increasing the Content of Nitrogen, Carbon, and Chlorophyll. *Microorganisms* **2021**, *9*, 1714. https://doi.org/10.3390/microorganisms9081714

Academic Editor: James F. White

Received: 3 July 2021
Accepted: 9 August 2021
Published: 11 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Endophytic microbes are an important component of Plant Growth Promoting Bacteria (PGPB), are ubiquitously present in all plants and do not cause disease symptoms [1]. They are often adapted to extreme habitats, developing strategies to induce stress responses in the host plants in which they reside [2]. The endophytes, being able to colonize the internal plant tissues, communicate and interact with the host plant more efficiently than

rhizospheric colonizers. In addition, once they have established within the tissues of the host plants, they are no longer subjected to the variation of changing soil conditions [1,2]. The endophytic PGPB are currently attracting a good deal of research interest due to their ability to help their host plants in getting increased amounts of limiting plant nutrients, including nitrogen, iron, and phosphorus, and to protect them against abiotic and biotic stresses, thus promoting their growth [3]. Soil salinity, more than any other factor, imposes a major constraint to the yield of agricultural crops, such as maize, rice, and sugarcane. It affects plant growth by increasing stress factors, such as ethylene, Na^+, and reactive oxygen species (ROS), which are harmful to the plant physiology, thus leading to growth defects [4,5].

To protect itself from oxidative damage, the plant produces anti-oxidative enzymes, which scavenge excessive ROS generated during salinity stress [6]. These enzymes include superoxide dismutase (SOD), catalase (CAT), peroxidase (POX), and ascorbate peroxidase (APX). Under stress conditions, the enhanced activities of anti-oxidative enzymes have been reported as a tolerance mechanism in plants [7]. Endophytic PGPB can assist plant growth under salinity stress through the stimulation of growth and the action of 1-aminocyclopropane-1-carboxylate (ACC) deaminase. ACC deaminase, encoded by the *acdS* gene, is responsible for the breakdown of ACC, the direct precursor of ethylene in all higher plants, into ammonia and α-ketobutyrate, used by bacteria as nitrogen and carbon sources [8,9]. The increase of ACC deaminase activity by overexpression of the corresponding gene has been reported to relieve environmental stresses of the host plants as compared to plants inoculated with the wild-type strains or the uninoculated ones [10].

Rice, a staple food and primary crop grown worldwide, feeds almost half of the world's population, particularly in Asia, Africa and Latin America. There are two distinct types of domesticated rice: *O. sativa* L. (Asian rice) and *O. glaberrima* Steud (African rice). African rice was domesticated about 2000–3000 years ago from the wild ancestor *O. barthii* by people living in the floodplains of the upper Inland Delta of the Niger River (Mali) [11]. Morpho-agronomic and molecular characterization of *O. glaberrima* accessions confirmed the existence of a relatively high reservoir of genetic diversity in this area [12]. Although there are farmers in West Africa still cultivating *O. glaberrima*, due to its low performance caused by the lodging and shattering, it is being massively replaced by Asian rice varieties. However, with a minimal intervention of humans and without the addition of nitrogen fertilizers, *O. glaberrima* acquired and developed tolerance and resistance against the majority of the biotic and abiotic stresses affecting rice cultivation in Sub-Saharan Africa [13,14]. Previous studies showed that rice plants (*O. sativa* L. cv. Baldo) adapted to temperate climates and were endophytically colonized by bacterial strains having plant growth-promoting (PGP) traits, which had positive changes in chlorophyll and nitrogen content, and shoot dry weight compared to uninoculated control plants [15], suggesting that auxin production and nitrogen-fixing abilities could be relevant traits of PGPB with a potential application in rice.

In this study, for the first time, the African rice *O. glaberrima* was selected to isolate cultivable bacterial endophytes. Considering that the root-associated microbes are affected by geographical location, soil source, host genotype, cultivation practice, and environmental stresses [16,17], we suppose that the population of bacterial endophytes hosted by *O. glaberrima* rice cultivated in the Niger River valley should be different from the one hosted by *O. sativa*, a tropical cereal extensively bred in Europe, and subjected to long-term agrochemical applications, and was selected for growth in a temperate climate with cold winter. Therefore, the main objective of this study was to establish a collection of relevant endophytes. To achieve this goal, we planned to: (1) isolate cultivable endophytic bacteria from the tissues of *O. glaberrima* plants (primary host); (2) characterize the selected endophytes for their plant growth-promoting (PGP) traits such as IAA-production, nitrogen-fixation and presence of the *acdS* gene, coding for ACC-deaminase enzyme; (3) analyse the ability of the N-fixing bacteria to colonize *O. sativa* plants (secondary host) and estimate the effect

of their inoculation on growth parameters of the host plants; 4) evaluate the response of inoculated plants to salt stress.

2. Materials and Methods

2.1. Isolation of Endophytic Bacteria from O. glaberrima Rice Tissues

O. glaberrima plants were kindly provided by Fousseyni Cisse and Amadou Sidibé from Institut d'Economie Rurale, Bamako. The Inner Niger Delta suffers from degraded soil quality and severe environmental stresses. The most important stress is the variation in the water regime, which is caused by both climate variability and upstream use of water. Endophytic bacteria were isolated from whole, healthy *O. glaberrima* rice plants grown up to 10 days in the Inner Niger Delta (Mali).

Plants were surface sterilized as follows: 1 min in 70% ethanol, 1 min in 5% sodium hypochlorite (NaClO), 30 s in 70% ethanol, and washing several times with sterile distilled water [15]. When sterilized plants were placed on LB agar plates and incubated at 30 °C for 3 days, no microbial growth was observed around the plant tissues. Sterilized samples were then ground with 5 mL of 1× PBS buffer using a sterile mortar and pestle. For the isolation of the endophytic bacteria, the tissue extracts were serially diluted in 1× PBS solution and plated onto R-2A plates (Sigma-Aldrich, St. Louis, MO, USA). The plates were incubated for up to 5 days at 30 °C to isolate fast-growing endophytes. Single colonies with different sizes, shapes and colors were selected and then streaked onto new R-2A plates. This procedure was repeated several times in order to isolate purified colonies. All purified colonies were grown in R-2A broth (HiMedia, VWR, Radnor, PA, USA) at 30 °C for 2 days. The bacterial cultures in exponential growth phase (OD_{600} ± 0.7) were stored at −80 °C in a solution containing 16% (*v*/*v*) dimethyl sulfoxide (DMSO) and 10% (*v*/*v*) glycerol [18].

2.2. Bacterial Identification by 16S rRNA Sequence Analysis

2.2.1. DNA Extraction

The genomic DNA from each strain's culture was isolated using the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA), following the manufacturer's instructions. The quality of DNA samples was analysed by 1% agarose gel electrophoresis. For DNA quantification 2 µL of the extracted genomic DNA were analysed by the NanoDrop 2000 spectrophotometer (Thermo Scientific, Walthman, MA, USA).

2.2.2. PCR Amplification of the 16rRNA Gene Fragment

To amplify the 16S rRNA gene fragment (1.4 Kb), universal primer pairs were used: 8-27F primer (5′-AGA GTT TGA TCC TGG CTC AG-3′), and 1510-1492R primer (5′-ACG GCT ACC TTG TTA CGA CTT-3′). The amplification reaction (25 µL) contained: 2.5 µL of 10x DreamTaq Buffer (Thermo Scientific, Walthman, MA, USA), 2 mM $MgCl_2$, 100 ng of the genomic DNA, 200 nM of each primer, 1 mM of each dNTP, and 0.03 U of DreamTaq Hot Start DNA Polymerase (Thermo Scientific, Walthman, MA, USA). A PCR mix without DNA was included as a negative control. The PCR reaction conditions were as follows: 3 min at 95 °C, 30 cycles of denaturation at 95 °C for 30 s and annealing at 52 °C for 30 s, extension at 72 °C for 1 min, and a final extension at 72 °C for 10 min [18].

2.2.3. Cloning and Sequencing of the 16S Fragment

The amplified 16S rRNA gene fragments were introduced into a linearized pHTP1 vector and cloned into competent cells of *E. coli* DH5α using the NZYEasy Cloning and Expression Kit I (NZYTech, Lisboa, Portugal). The 16S rRNA gene inserts were prepared by PCR using primers with 16bp overhangs included on 5′-ends of both forward and reverse primers, in order to provide vector-complementary single-strands terminals. The following primers were designed: 16S-NZYf (5′-T CAG CAA GGG CTG AGG CCT AGA GTT TGA TCA TGG CTC AG-3′) and 16S-NZYr (5′-T CAG CGG AAG CTG AGG GGT TAC CTT GTT ACG ACT T-3′). The preparation of DNA inserts was carried out with a PCR amplification reaction (25 µL) containing 0.5 µL of 16S rRNA gene PCR products, above obtained, 200 nM

of the 16S-NZY primer pairs, 2.5 µL of 10x Buffer, 2 mM $MgCl_2$, 1 mM of each dNTP, 0.03 U of DreamTaq Hot Start DNA Polymerase (Thermo Scientific, Walthman, MA, USA). The PCR reaction conditions were as follows: 3 min at 95 °C, 30 cycles of denaturation at 95 °C for 30 s and annealing at 68 °C for 30 s, extension at 72 °C for 1 min, and a final extension at 72 °C for 10 min. The inserts were purified using the QIAquick PCR Purification Kit (QIAGEN, Hilden, Germany) and transferred into pHTP1 vector.

A ligase-independent cloning reaction was prepared following the manufacturer's instruction, using a vector to insert molar ratio of 1 to 5. The cloning product was used to transform competent cells of *E. coli* DH5α. The recombinant clones were identified by colony PCR reactions. For the amplification of the insert, the 16S-NZYf and 16S-NZYr primers pair and the PCR reaction conditions described above were used. The PCR products were purified, and their nucleotide sequences determined through DNA fragments sequencing by fluorescent chain termination and capillary electrophoresis method (Eurofins, Ghent, Belgium). The obtained sequences were first trimmed to eliminate the low-quality ends and then compared with those deposited in GenBank database (www.ncbi.nlm.nih.gov/BLAST/, accessed on 25 March 2020) using the BLASTN algorithm (BLAST Local Alignment Search Tolls).

2.2.4. Phylogenetic Characterization

The 16S rRNA gene sequences were first aligned by using ClustalW and then analysed by MEGA6 software [19] to infer the molecular phylogeny by the maximum likelihood method based on Tamura–Nei model [20]. The robustness of the phylogenetic tree was evaluated by bootstrap analysis of 1000 pseudo-replications. The scale bar on the rooted tree indicates 0.05 substitutions per nucleotide position.

2.3. Genome Sequencing and Bioinformatic Analysis

DNA was extracted from overnight grown cultures in LB medium by using PowerSoil DNA Isolation Kit (Qiagen, Hilden, Germany) After gel electrophoresis and fluorometric quantification (Qubit, Thermo Fisher Scientific, Waltham, MA, USA), genomic DNA was fragmented with g-TUBE (Covaris Inc., Woburn, MA, USA) to average 15 Kbp size and used for library preparation using Pacific Biosciences SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences, Menlo Park, CA, USA). Samples were sequenced with the Sequel Sequencing Kit 3.0 in a Sequel apparatus (Pacific Biosciences, Menlo Park, CA, USA) by SMRT technology [21]. Obtained reads were analysed using SMRT Link software ver. 8.0.0.80529 (Pacific Biosciences, Menlo Park, CA, USA) producing *oriC*-oriented assembly by running the "microbial multiplexing" pipeline with default options. Annotation and the COG assignment were performed by Prokka 1.4.0 [22] from the Galaxy webserver Orione (https://orione.crs4.it/, accessed on 1 April 2021) [23]. Taxonomic identification of strains from genome sequence was performed using ContEst16S [24] and Type (Strain) Genome Server [25]. Presence of CRISPR arrays was checked by CRISPRFinder [26]. Secondary metabolites gene clusters have been identified by AntiSMASH [27]. The presence of gene modules for chemotaxis, flagellar biosynthesis and secretion systems were inspected by performing KAAS functional annotation [28]. The presence of Type III secretion systems was additionally anayzed by T3Sepp [29]. Sequences are deposited in NCBI database under Bioproject (PRJNA670042).

2.4. Screening of Plant Growth-Promoting (PGP) Activities

2.4.1. Indole-3-Acetic Acid (IAA) Production

Salkowski colorimetric assay. The IAA production into bacterial cultures was analysed by using the Salkowski reagent [30] as described below. Bacterial culture was grown overnight in LB medium containing 100 µM L-tryptophan at 30 °C. Liquid culture was centrifuged at 12,000 rpm for 5 min to collect the bacterial cells. The culture supernatant was then mixed with the Salkowski reagent (35% $HClO_4$ 49 mL and 0.5 M $FeCl_3$ 1 mL) in a 1 to 1 ratio and incubated in the dark at room temperature. After 30 min of incuba-

tion, the absorbance of the mixtures was estimated at 530 nm. The bacterial cells were dried and weighed for data normalization. The concentration of IAA from each culture medium was calculated from a pure IAA standard curve. Data are the mean ±SD of at least five independent bacterial cultures each grown at different times and from different bacterial colonies.

2.4.2. HPLC—Tandem Mass Spectrometry (HPLC-MS/MS)

The high-pressure liquid chromatography (HPLC) column was an Ultra AQ C18 (3 µm pore size, 100 × 2.1 mm), operating at 250 µL/min flow rate and thermostated at 35 °C. The mobile phases were 5 mM ammonium acetate in LC-MS grade H_2O containing 0.2% formic acid (mobile phase A) and acetonitrile containing 0.2% formic acid (mobile phase B). For gradient elution, the percentage of mobile phase B was linearly increased from 0% to 40% in 11 min, then to 90% in 2 min; after 7 min at 90% mobile phase B, the initial composition was restored and the system equilibrated for 10 min. The injection volume was 10 µL.

The HPLC was a Series 200 pump with a Series 200 autosampler (Perkin Elmer, Milano, Italy) coupled to a 4000 Q Trap mass spectrometer (Sciex, Old Connecticut Path Framingham, MA, USA), equipped with a Turbo V IonSpray interface. The interface and MS parameters were optimized by infusion of a standard solution of IAA (Sigma-Aldrich, St. Louis, MO, USA). The settings used were the following: curtain gas 25, nebulizer gas 70, heater gas 68, temperature 550 °C, ESI spray voltage, 5.2 kV; declustering potential (DP) 40 V, entrance potential (EP) 10 V. Nitrogen was used as curtain, nebulizer, and heater gas. Acquisition was done in multiple reaction monitoring (MRM) mode in positive ion polarity, recording two MRM transitions for IAA and its isotopologue, indole-2,4,5,6,7-d5-3-acetic-2,2-d2 acid (d7-IAA; purchased from CDN Isotopes, Pointe-Claire, Canada): the protonated ions of IAA *(m/z* 176.1) and d7-IAA (*m/z* 183.1) were fragmented in the collision cell of the MS and the product ions at *m/z* 130 (collision energy, CE, 23 V) and *m/z* 103 (CE 36 V) for IAA, and *m/z* 136 (CE 23 V) and *m/z* 109 (CE 40 V) for d7-IAA were recorded. For HPLC-MS/MS analyses, 1 µL of a solution of d7-IAA (25 ng/µL in methanol) was added to 100 µL of sample and then injected in the HPLC-MS/MS instrument. Peak identification was accepted only if the relative retention time of IAA was consistent with that of the internal standard d7-IAA (±0.5%) and if the ratio of the signals of the two transitions were in agreement with that recorded for the standard. For quantitative purposes, ratio of the peak areas (PAR, peak area ratio) of the signals corresponding to the most intense MRM transition for IAA (*m/z* 176 → *m/z* 130) and for d7-IAA (*m/z* 183 → *m/z* 136) was measured for each sample. The sample concentration was calculated using a six-point linear calibration curve prepared adding scalar amounts of IAA to several aliquots of blank culture medium to mimic actual samples in the concentration range from 0.025 ng/mL to 0.5 ng/mL. The calibration points were processed in parallel to the samples, following the same procedure. For each sample three 100 µL aliquots were measured in triplicate.

2.5. PCR Amplification of the Nitrogenase Iron Protein (nifH) and ACC Deaminase (acdS) Genes Fragments

The PCR amplification of the *nifH* and *acdS* genes was performed to identify nitrogen-fixing (diazotrophic) and ACC-deaminase-producing (halotolerant) bacteria. The primer pairs used were: 19F (5′-GCI WTY TAY GGI AAR GGI GG-3′) and 407R (5′-AAI CCR CCR CAI ACI ACR TC-3), which amplified a 388 bp fragment of the *nifH* gene [31] acdSF3 (5′-ATCGGCGGCATCCAGWSNAAYCANAC -3′) and acdSR3 (5′-GTGCATCGACTTGCCCTC RTANACNGGRT-3), which amplified a segment of 683 bp of the *acdS* gene [32]. The amplification reactions (25 µL) for the screening of the *nifH* and *acdS* genes contained: 2.5 µL of 10x DreamTaq Buffer (Thermo Scientific, Walthman, MA, USA), 2 mM $MgCl_2$, 100 ng of the genomic DNA, 200 nM (*nifH)* and 400 nM (*acdS*) of each primer, 1 mM of each dNTP, and 0.03 U of DreamTaq Hot Start DNA Polymerase (Thermo Scientific, Walthman, MA, USA). The thermocycling conditions for *nifH* and *acdS* genes were as follows: 30 sec at 94 °C, 35 cycles of denaturation at 95 °C followed by annealing at 50 °C (*nifH*) and 62 °C (*acdS*)

for 1 min, extension at 72 °C for 1 min, and a final extension at 72 °C for 10 min. Sterile Milli-Q water, instead of genomic DNA, was used to set up control reaction. As previously proposed by Zhengyi et al. [32], the DNA of *Herbaspirillum seropedicae* z67 (HSz67) was used as positive control for *acdS* gene fragments, while the DNA of *Sinorhizobium meliloti* 1021 was used as positive control for *nifH* gene fragments. The DNA of the *Escherichia coli* MG1655 was used as negative control for both genes.

2.6. Plant Growth and Inoculation Methods

'Baldo' rice has been selected to growth in Northern Italy (Piemonte and Lombardia regions) characterized by environmental conditions very different from the ones typical of West Africa. Indeed, the average annual temperature in Mali is around 27 °C, with few variations in different seasons, while in the regions of northern Italy temperatures fluctuate between 0 °C in winter and an average of 20 °C in summer. *O. sativa* L. cv. Baldo seeds were kindly provided by Stefano Monaco and Elisa Zampieri from CREA-CI, Research Centre for Cereal and Industrial Crops, Vercelli, Italy. Dehulled seeds of 'Baldo' and *O. glaberrima* RAM133 (Mali, West Africa) were surface sterilized as described by Defez et al. [18] with the following changes: 70% ethanol for 7 min and 5% sodium hypochlorite solution containing 0.1% tween 20. Seeds were then washed several times with sterilized distilled water, positioned onto the surface of 0.8% water-agar plates and incubated at 21 °C in the dark for germination [18].

After 5 days, the roots of the germinated seeds were cut (0.5 cm from the bottom) with a sterile bistoury and incubated in Petri dishes with 50 mL of 1× PBS solution containing each strain to a final concentration of 10^6 cells mL^{-1} for 4 h at room temperature [15]. Seeds incubated in 1× PBS were used as control. Infected seeds were then transferred into plastic pots containing sand (2–1.2 mm granule size) and perlite (3–4 mm granule size) soil in 1:1 ratio, or into hydroponic units (plastic tubes of 15 cm in length and 5 cm in diameter) containing Kimura B nitrogen-free medium reported in [33] with some modifications.

The macronutrients were the following: KH_2PO_4 (0.2 mM), KNO_3 (0.01 mM), K_2SO_4 (0.1 mM), $CaSO_4$ (0.4 mM), $MgSO_4 \cdot 7H_2O$ (0.5 mM). The micronutrients were the following: Fe-EDTA (0.11 mM), $MnSO_4$ (1.8 µM), H_3BO_3 (46 µM), $ZnSO_4 \cdot 7H_2O$ (0.3 µM), and $CuSO_4 \cdot 5H_2O$ (0.3 µM). Each planting unit was kept in the growth chamber under long daylight (16 h), 19–23 °C temperature and 75% relative humidity.

2.7. Re-Isolation of Diazotrophic Bacteria from Inoculated Plants

The colonization of *O. sativa* cv. Baldo and *O. glaberrima* RAM133 plants by the selected endophytes was assessed at 14 days after inoculation (DAI) for plants grown in sand-perlite soil. The surface of the whole plants was sterilized using the following procedure: (i) 1 min in 70% EtOH; (ii) washing several times with sterile distilled water; (iii) incubation for 5 min in 5% sodium hypochlorite solution containing 0.1% tween 20; (iv) washing several times with sterile distilled water. The tissues of the whole plant were then homogenized in 5 mL of 1× PBS using a sterile mortar and pestle. Dilutions of the homogenates (1:100 for the extract of Baldo and 1:1000 for RAM133) were spread onto 1.5% LB agar plates containing the specific antibiotics (60 µg mL^{-1} ampicillin for *P. diazotrophicus* BDA59-3; 25 µg mL^{-1} carbenicillin for *Ko. pseudosacchari* BDA62-3, *Kl. pasteurii* BDA134-6 and *Ko. oryzendophytica* BDA137-1; 25 µg mL^{-1} penicillin G for *E. sacchari* BDA86-11; 50 µg mL^{-1} novobiocin for *Enterobacter* sp. BDAM41-2) and incubated at 30 °C. Specific antibiotics were selected after analysing the resistance of the selected strains to different antibiotics (Table S1). After 24 h, the CFU number was counted. The identity of re-isolated endophytes was verified by their antibiogram patterns. To confirm that the sterilization process described above was successful, sterilized plant tissues were placed on 1.5% LB agar plates and incubated at 30 °C for 3 days. After the incubation period, no microbial growth was observed around the plant tissues. Data for plants grown in sand-perlite soil and hydroponic conditions are the mean ± SD of at least five and eight independent replicates, respectively, each carried out at different times.

2.8. Strains Tagging with FPs

Plasmids pMP4655 (harbouring the EGFP marker) [34,35] were purified from *Escherichia coli* DH5α host cells using the PureLink™ HiPure Plasmid Filter Midiprep Kit (Thermo Scientific, Walthman, MA, USA) according to the manufacturer's instructions. The transformation of the strains BDA62-3 and BDA134-6 was carried out by electroporation method, while that of the strain BDA59-3 was done by triparental mating. Protocols for the electrocompetent cells preparation, the electroporation and the triparental mating were those reported in [15].

2.9. Confocal Laser Scanning Microscopy Detection

To elucidate the colonization process of in vitro rice plants by BDA62-3, BDA134-6 and BDA59-3 strains, one plant per bacteria treatment was analysed each day from 15 days after bacterization. At least ten different whole roots from each rice plant were mounted with distilled water for confocal visualization. Confocal laser scanning microscopy (CLSM) stacks were obtained using an Axioskop 2 MOT microscope (Carl Zeiss Inc., Göschwitzer Straße, Jena GmbH, Germany) equipped with a krypton and an argon laser and supported by the Laser Scanning System LSM5 PASCAL software (Carl Zeiss Inc., Göschwitzer Straße, Jena GmbH, Germany). For microscope data analysis, the Zeiss LSM Image Browser version 4.0 program (Carl Zeiss Inc., Göschwitzer Straße, Jena GmbH, Germany) was used. Bacterial colonization of rice roots and shoots was analyzed from 3-D confocal data stacks. Projections from adjacent confocal optical sections were made for composing images shown in this study. Final figures were processed with Photoshop 4.0 software (Adobe Systems Inc., San Jose, CA, USA).

2.10. Nitrogenase Activity by Acetylene Reduction Assay (ARA)

2.10.1. Bacterial Cultures

Strains selected as positives in the PCR analysis of *nifH* gene were grown in 1x M9 minimal ammonium chloride medium [36] for 48 h at 30 °C on a rotary shaker (200 rpm). The cell pellets (2 mL) were suspended in the same medium without nitrogen source, as described by Defez et al. [36] and transferred into glass tubes airtight with a serum cap. To assess nitrogenase activity, we used the acetylene reduction assay (ARA) [37]. Purified acetylene was injected into each tube at approximately 10% of headspace volume under a hypoxic atmosphere (2% O2). After 18 h of incubation at 30 °C, a 1 mL sample from each bottle was injected into a gas chromatograph GC Clarus®580 (Perkin-Elmer, Walthan, MA, USA) to evaluate the amount of the produced ethylene (C_2H_4) after the addition of acetylene. A parallel set of C_2H_4 standards was injected into the GC to calibrate its sensitivity, linearity, and a concentration of C_2H_4 produced in samples. The GC used was equipped with a hydrogen flame detector (FID) and a TG-IBOND Alumina (Na_2SO_4 deactivate) column (30 m × 0.53 mm × 10 µm) (Thermo Scientific, Walthman, MA, USA). The flow rate of the carrier (helium) was 48 cm s^{-1}, and the oven programmer was 130 °C isocratic for 3 min. Inoculated tubes without injected acetylene were used as a negative control. Data are expressed as nmol ethylene mg $cells^{-1}$ min^{-1} and are the mean ± SD of at least five independent bacterial cultures, each grown at different times and from different bacterial colonies.

2.10.2. Inoculated Plants

Two-week-old rice plants (*O. sativa* cv. Baldo and *O. glaberrima* RAM133), grown as previously described, were carefully removed from the pots and the roots rinsed with water. The plants were then transferred into 20 mL glass tubes, incubated in a greenhouse for 20 h and analysed for the ARA assay as described by Defez et al. [18]. The amount of ethylene produced was measured by gas chromatography as described above. Data for plants grown in sand-perlite soil and hydroponic conditions are expressed as nmol ethylene $plant^{-1}$ min^{-1} and are the mean ± SD of at least six and eight independent replicates, respectively, each carried out at different times.

2.11. $^{15}N_2$ Incorporation

Two-week-old rice plants (*O. sativa* cv. Baldo) non-inoculated and inoculated with the strains BDA59-3, BDA62-3 and BDA134-6 were removed from the pots and washed several times with sterilized distilled water. Non-inoculated plants were used as references.

The plants were placed into 20-mL bottles containing 2 mL of minimal medium free of nitrogen sources and sealed with a rubber serum stopper. The bottles containing an atmosphere of 15% (vol/vol) $^{15}N_2$ (99.9 atom%), 20% (vol/vol) N_2, and 10% (vol/vol) O_2 were incubated in greenhouse for 24 h in the dark. The plants were then taken from the bottles, freeze-dried and ground in a mortar. The $^{15}N/^{14}N$ ratios were measured using an isotope mass spectrometer (Elementar Gmbh, Isoprime Ltd., Langenselbold, Germany) after total combustion in an elemental analyser (VARIO CUBE, Isoprime Limited, Langenselbold, Germany). To analyse the samples the amount weighed was 3.5 mg. The samples were analysed in triplicate and reported as an average.

Due to the small variations recorded, the data is not reported as atom% in excess with respect to the non-inoculated samples but, according to the IUPAC protocol, they are denoted in delta in relation to the international standard Air (atmospheric N_2 with 0.366 atom%) for $\delta^{15}N$, according to the following general equation: δi E = (i RSA–i RREF)/ i RREF [38], where *i* is the mass number of the heavier isotope of element E (^{15}N); *RSA* is the respective isotope ratio of a sample (such as number of ^{15}N atoms/number of ^{14}N atoms or as approximation $^{15}N/^{14}N$); *RREF* is the respective isotope ratio of internationally recognized reference material. The delta values are multiplied by 1000 and expressed in units "per mil" (‰). The isotopic values were calculated against working in-house standards, which were themselves calibrated against international reference materials: L-glutamic acid USGS 40 with $\delta^{15}N = -4.52$‰ (U.S. Geological Survey, Reston, VA, USA) and potassium nitrate IAEA-NO3 ($\delta^{15}N = +4.7$‰) from IAEA (International Atomic Energy Agency, Vienna, Austria). The uncertainty of measurements, calculated as standard reproducibility obtained from the analysis of the same sample over time and multiplied for the coverage factor 2, was <0.3‰ for $\delta^{15}N$.

2.12. Salt Stress

To apply salt stress, three-week-old plants were carefully removed from the pots and the roots rinsed with water. Plants were then transferred into 20 mL glass tubes airtight with serum cap containing 2 mL of minimal medium supplemented with NaCl at a final concentration of 0.3 M and free of nitrogen sources. Plants were then incubated in greenhouse for 20 h and analysed for the ARA test as described by Defez et al. [18]. After the ARA test, plant tissues (roots and leaves) were immediately collected, frozen in liquid nitrogen and stored at −80 °C for further analyses.

2.13. Antioxidant Enzymes Assays

The soluble leaf proteins were extracted and quantified as described in Defez et al. [18]. The activities of all enzymes [ascorbate peroxidase (APX), catalase (CAT), peroxidase (POX) and superoxide dismutase (SOD)] were measured spectrophotometrically at 25 °C within the linear region for both time and enzyme concentration as reported in Bianco and Defez [39].

2.14. Chlorophyll Content Measurement

At 14 and 40 DAI detached leaves were incubated at room temperature in a 2 mL tube containing 1.5 mL 80% acetone solution for at least 24 h and then clarified by centrifugation for 5 min at 15,000× *g*. The absorbance of the supernatant was measured at wavelengths 645, 646, and 663 nm (A_{645}, A_{646}, A_{663}) [40] with a DU 800 UV/Visible spectrophotomer (Beckman Coulter, Brea, CA, US). Chlorophyll concentration was estimated following the Arnon's equations [41] as follows:

Chlorophyll a (μg/mL) = 12.7 (A_{663}) − 2.69 (A_{645})

Chlorophyll b (μg/mL) = 22.9 (A_{645}) − 4.68 (A_{663})

Total chlorophyll (μg/mL) = 20.2 (A_{645}) + 8.02 (A_{663})

Data are the means ± SD of at least five independent biological replicates each carried out at different times.

2.15. Total Carbon and Nitrogen Analysis

Total organic carbon and nitrogen content from *O. sativa* leaves at 14 and 40 DAI were determined with dried material by using an elemental analyser (CKIC 5E-CHN 2200, Emme3, Lainate, Milano, Italy) at 950 °C, according to the European standard method UNI EN 15407:2011. Data are the means ±SD of at least five independent biological replicates each carried out at different times.

2.16. Statistical Analysis

Five biological replicates were carried out for the measurement of IAA production, endophytes re-isolation from inoculated plants, nitrogenase activity for bacterial cultures, nitrogen content, carbon content, and chlorophyll content. Six biological replicates have been carried out for the activity of nitrogenase and anti-oxidative enzymes in salt-stressed inoculated plants. The evaluation of nitrogenase activity in inoculated plants under hydroponic conditions was performed on eight biological replicates. Statistical analysis of nitrogenase activity, physiological parameters (N, C, and Chl) and anti-oxidative enzymes activity was performed by applying the one-way analysis of variance (ANOVA) and Tukey's HSD post-hoc test. All data are approximately normally distributed. The VassarStats ANOVA program available at http://vassarstats.net/index.html (accessed on 8 June 2020) was used. The remaining data were subjected to Student's *t*-test. All results were considered statistically significant when $p \leq 0.05$.

3. Results

3.1. Identification of Endophytic Bacteria

We here report the isolation of 69 culturable endophytic bacteria from internal tissues of African rice plants (*O. glaberrima* L.). All bacterial isolates were identified by comparative sequence analysis of 16S rRNA gene partial sequences. For all isolates, a 99–100% similarity with already identified species in the GenBank database was found (Table S2). The isolates belonged to four phyla: Proteobacteria (79%, Aeromonadales, Burkholderiales, Enterobacteriales, Pseudomonadales and Xantomonadales orders), Firmicutes (7%, Bacillales), Actinobacteria (10%, Actinomycetales), and Bacteroides (4%, Sphingobacteriales and Flavobacteriales) (Figure 1 and Table S2).

The evolutionary history was inferred by using the maximum likelihood method based on the Tamura–Nei model with 1000 bootstrap replications. The percentage of trees in which the associated taxa clustered together is shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 69 nucleotide sequences. Evolutionary analyses were conducted in MEGA6.

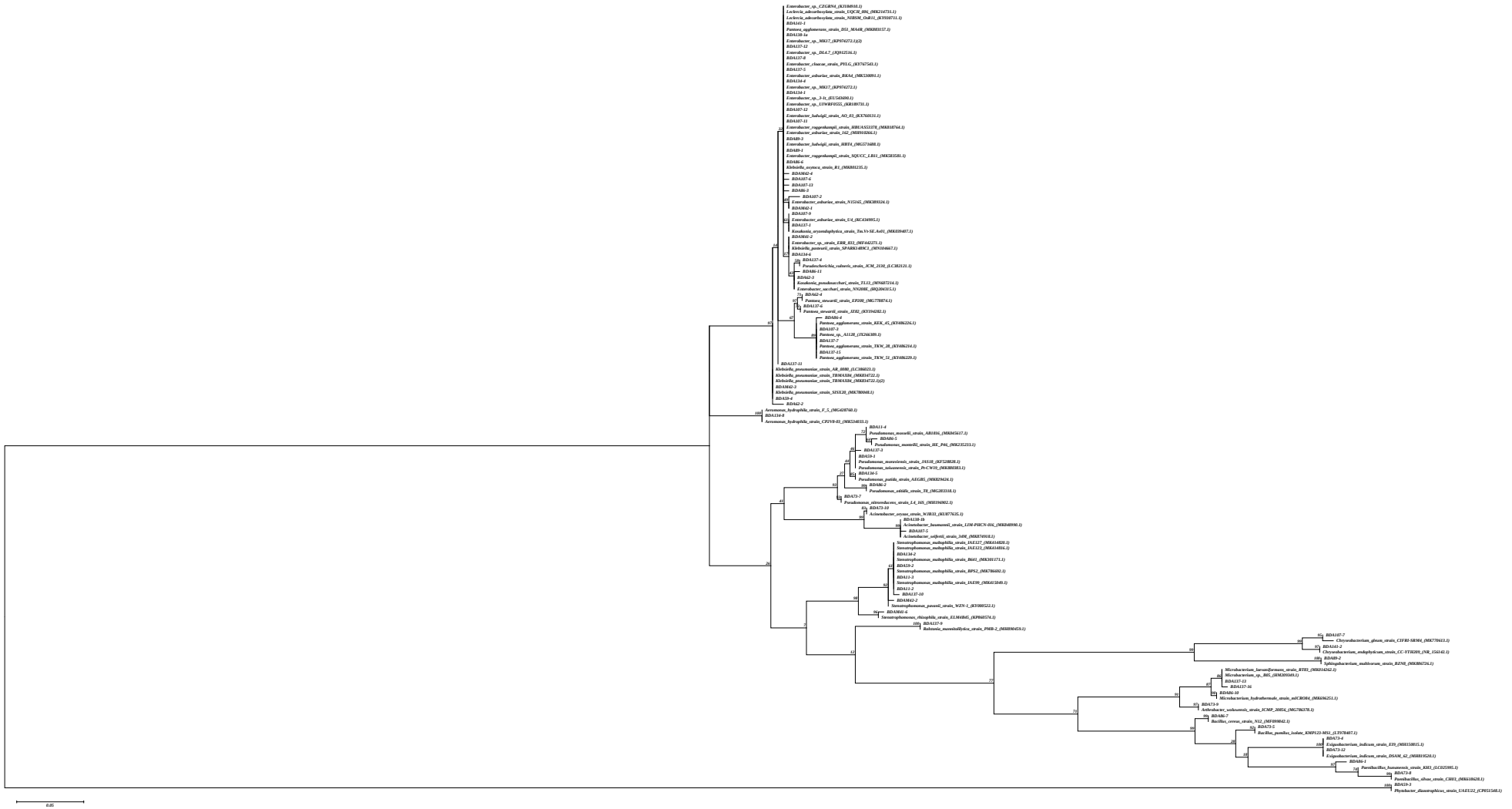


Figure 1. Phylogenetic tree based on the 16S rRNA gene sequences corresponding to endophytic strains inhabiting rice tissues. The sequences for closely related strains were recovered from GenBank and included in the tree. The accession numbers of these strains are shown in parenthesis.

3.2. PGP-Traits of Bacterial Endophytes

All isolates were examined for the IAA production and for the presence of genes involved in N-fixation (*nifH*) and in the reduction of ethylene level inside the plant (*acdS*). The results obtained in these analyses are reported in Table S2. When the L-tryptophan was added to the growth medium the 67% of the bacterial strains produced IAA. The IAA produced by the three N-fixing strains *Kl. pasteurii* BDA134-6, *Ko. pseudosacchari* BDA62-3and *P. diazotrophicus* BDA59-3 was quantified by using liquid chromatography coupled to high-tandem mass spectrometry. High signal-to-noise ratios and, therefore, enhanced sensitivity were obtained by applying this procedure. The deuterated IAA (d7-IAA) molecule was used as an internal standard for quantitative purposes. Appreciable IAA level was measured only for the strain *Kl. pasteurii* BDA134-6 (0.44 ± 0.04 ng mL^{-1}). Forthe strains *Ko. pseudosacchari* BDA62-3 and *P. diazotrophicus* BDA59-3 IAA level (<0.025 ng mL^{-1}) below the threshold of the instrument was observed. For the strain *P. diazotrophicus* BDA59-3 higher value was detected with the spectrofotometric method as compared to mass spectrometry. This discrepancy could result from the well-known ability of the genus *Citrobacter* (some of them are now classified as *P. diazotrophicus*) [42,43] to produce IAA [44] and from the reaction of the total indole compounds with the Salkowski reagent, which alters the IAA evaluation.

The 388 bp amplicon expected for the *nifH* gene was observed for six (9%) bacterial strains. Positive isolates were *Ko. oryzendophytica* BDA137-1, *Enterobacter* sp. BDAM41-2, *Kl. pasteurii* BDA134-6, *Enterobacter sacchari* BDA86-11, *Ko. pseudosacchari* BDA62-3, and *P. diazotrophicus* BDA59-3The diazotrophic nature of the six *nifH*-positive strains was confirmed by measuring the activity of the nitrogenase enzyme in the bacterial cultures (ARA test). The data reported in Table 1 showed that the strain *Kl. pasteurii* BDA134-6 was the best N-fixer. The strains *E. sacchari* BDA86-11, *Ko. pseudosacchari* BDA62-3 and *P. diazotrophicus* BDA59-3 showed lower activity than the one recorded for *Kl. pasteurii* BDA134-6. The lowest nitrogenase activity was measured for the strains *Ko. oryzendophytica* BDA137-1 and *Enterobacter* sp. BDAM41-2. Concerning the *acdS* gene, amplicons of the expected size (683 bp from *H. seropedicae* z67) were obtained for six isolates belonging to the genera *Enterobacter* (*Enterobacter* sp. BDAM41-2 and *E. roggenkampii* BDA86-6), *Stenotrophomonas* (*S. rhizophila* BDAM41-6), *Ralstonia* (*R. mannitolilytica* BDA137-9), and *Microbacterium* (*M. laevaniformans* BDA137-13 and *Microbacterium* sp. BDA137-16).

Table 1. Nitrogenase activity and IAA production in bacterial cultures of N-fixing endophytes isolated from *O. glaberrima.*

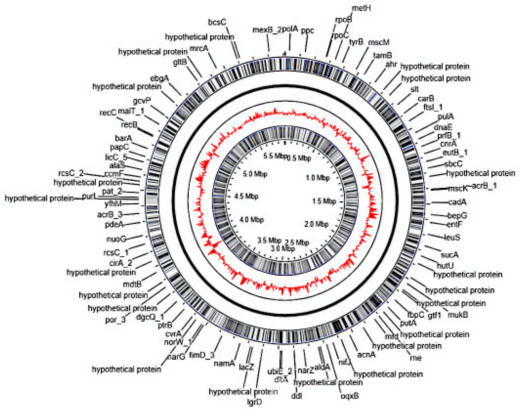
Strain	Nitrogenase Activity § (nmol Ethylene/mg Cells/min)	IAA (nmol/mg Cells) ξ
Ko. oryzendophytica BDA137-1	0.032 ± 0.004^{c}	8.7 ± 1.4^{c}
Enterobacter sp. *BDAM41-2*	0.025 ± 0.005^{c}	6.1 ± 0.3^{c}
Kl. pasteurii BDA134-6	0.160 ± 0.024^{a}	31 ± 2^{b}
E. sacchari BDA86-11	0.064 ± 0.008^{b}	7.4 ± 0.3^{c}
Ko pseudosacchari BDA62-3	0.049 ± 0.008^{b}	7.7 ± 0.5^{c}
P. diazotrophicus BDA59-3	0.044 ± 0.003^{b}	61 ± 10^{a}

§ Bacterial cells grown in M9 minimal ammonium chloride medium were suspended in the same medium without nitrogen source and incubated under a hypoxic atmosphere (2% O2) and 10% acetylene. After 18 h of incubation at 30 °C the amount of the ethylene produced was measured by using a gas chromatograph. ξ Bacterial cells were grown overnight in LB medium containing 100 µM L-tryptophan at 30 °C. The culture supernatant was then mixed with the Salkowski reagent and incubated for 30 min in the dark at room temperature. The absorbance of the mixtures was then estimated at 530 nm. The bacterial cells were dried and weighed for data normalization. The values reported in the table are the mean ± SD of at least five independent bacterial cultures each grown at different times and from different bacterial colonies. Means with different letters are significantly different at 5% level of confidence.

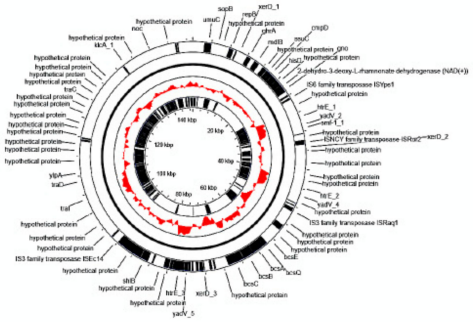
3.3. Genome Sequences of Kl. pasteurii BDA134-6, Ko. pseudosacchari BDA62-3 and P. diazotrophicus BDA59-3 Confirm the Presence of the Full Nitrogenase Genes Set

Single Molecule Real-Time (SMRT) genome sequencing of strains BDA62-3, BDA134-6 and BDA59-3 strains yielded complete genome sequences (Table S3). Genome-based taxonomic assignment indicated that BDA134-6 and BDA62-3 belong to the species *Kl. pasteurii* and *Ko. pseudosacchari*, confirming 16SrRNA gene sequences, while BDA59-3 belongs to the species *P. diazotrophicus* (Table S4), although the 16S rRNA gene sequences showed the best match with members of genus *Citrobacter* (data not shown). *Ko. pseudosacchari* BDA62-3 and *Kl. pasteurii* BDA134-6 harbour two replicons for a total genome size of ~5.0 Mbp and ~6.0 Mbp, respectively, while *P. diazotrophicus* BDA59-3 genome consists of a single chromosome ~5.3 Mbp in size (Figure 2). *Ko. pseudosacchari* BDA62-3 contains a chromosome ~4.9 Mbp in size and a plasmid, named pBDA62-3, of 93,051 bp. *Kl. pasteurii* BDA134-6 has a chromosome of ~5.9 Mbp and a plasmid, named pBDA134-6 of 149,270 bp. *P. diazotrophicus* BDA59-3 shows three CRISPR arrays, containing CRISPR-associated proteins similar to those found in *C. rodentium* ICC168 (74% sequence identity). *Kl. pasteurii* BDA134-6 harboured one CRISPR array on the chromosome; however, no *cas* genes were found. No CRISPR were found in *Ko. pseudosacchari* BDA62-3 genome. The methylated DNA motifs detected by SMRT sequencing technology were mostly due to m6A modification (6-methyl adenine) and concerned with the classical GATC motif [45] and 2 and 4 additional motifs in *Ko. pseudosacchari* BDA62-3 and *P. diazotrophicus* BDA59-3, respectively, some of them compatible with the presence of restriction-modification systems (e.g., CTCGAG corresponds to the *XhoI* recognition site, CTGAAG to *Eco57I*). For *Kl. pasteurii* BDA134-6 a m4C (4-methyl cytosine) modification motif was found (VSAGCTSS) present in more than 60% of sites (corresponding to the *AluI* recognition site AGCT). Genome annotation files are reported as supplemental material (Folder S1).

Kl. pasteurii BDA134-6

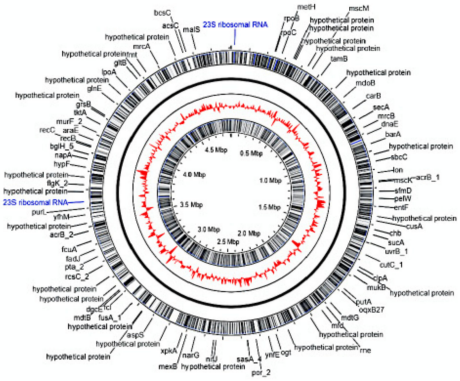


Chromosome: 5,912,274 bp

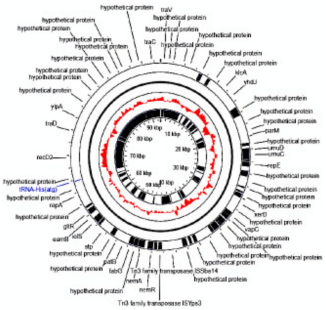


plasmid pBDA134-6: 149,270 bp

Ko. pseudosacchari sp.BDA62-3

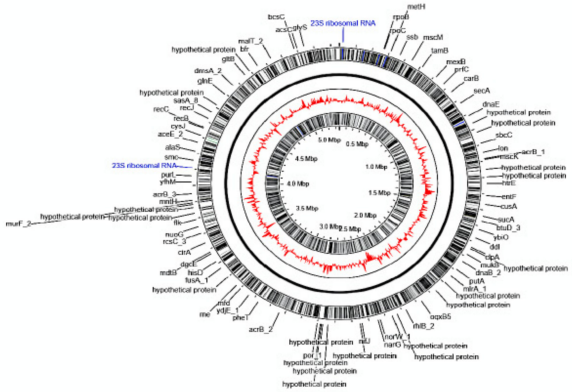


Chromosome: 4,909,996 bp



plasmid pBDA62-3: 93,051 bp

P. diazotrophicus BDA59-3



Chromosome: 5,349,552 bp

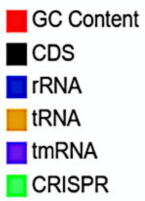


Figure 2. Circular maps of the genomes of *Kl. pasteurii* BDA134-6, *Ko pseudosacchari* BDA62-3 and *P. diazotrophicus* BDA59-3 showing the inner GC content and annotation. Maps were obtained from the CGView Server (http://cgview.ca/, https://doi.org/10.1093/bib/bbx081, accessed on 1 April 2021).

When looking for genes involved in nitrogen-fixation ability, shared within the three strains, the full set of genes involved in nitrogenase functioning was found, with 17, 12 and 18 orthologs assigned to nitrogenase complex functioning in *Ko. pseudosacchari* BDA62-3, *Kl. pasteurii* DA134-6, and *P. diazotrophicus* BDA59-3, respectively (Table 2 and Figure S1). In order to evaluate the presence of genetic determinants involved in the ability to associate with plants [46], we screened genome sequences of *Ko. pseudosacchari* BDA62-3, *Kl. pasteurii* BDA134-6 and *P. diazotrophicus* BDA59-3 for genes involved in chemotaxis, flagellum biosynthesis, Type III secretion systems, Type VI secretion system (T6SS), COG1609 and COG0667, which have been found enriched in genomes of plant-associated bacteria [46]. A type-III secretion system was retrieved for *Ko. pseudosacchari* BDA62-3 only. However, no effector proteins were identified. No T6SSs were found in *Ko. pseudosacchari* BDA62-3 and *P. diazotrophicus* BDA59-3, while in *Kl. pasteurii* BDA134-6 orthologs encoding for proteins of Hcp, VgrG, IcmF and ClpV were found, suggesting the presence of a complete set for T6SS in this strain. Both *Ko. pseudosacchari* BDA62-3 and *P. diazotrophicus* BDA59-3 displayed the presence of genes for chemotaxis and flagellum biosynthesis (Dataset S1), which include KO annotation lists and AntiSMASH results and Table S4. *Kl. pasteurii* BDA134-6 only showed *motB* (K02557) and *fliY* (K02424) genes, deserving further attention on chemotaxis and motility phenotypes. In *Ko. pseudosacchari* BDA62-3, *Kl. pasteurii* BDA134-6, and *P. diazotrophicus* BDA59-3 genes assigned to COG1609 were retrieved for 19, 28 and 21 genes, respectively, while for genes assigned to COG0677, only one gene, encoding an UDP-*N*-acetyl-D-mannosamine dehydrogenase, was found in all three genomes (Table S5). We also looked at the presence of gene clusters involved in the biosynthesis of secondary metabolites potentially involved in plant-microbe and microbe-microbe interactions (Folder S2). A bacteriocin gene cluster (located between 1,505,693–1,516,316 nt) and a nonribosomal peptide-synthetase (NRPS) gene cluster (located between 95,649–147,789 nt), showing 30% similarity with the gene cluster for the biosynthesis of the siderophore turnerbactin [47], was found in *P. diazotrophicus* BDA59-3. In *Ko. pseudosacchari* BDA62-3, a bacteriocin gene cluster (located between 18,651–29,274 nt) was found too, as well as two nonribosomal peptide-synthetase (NRPS) gene clusters (located between 1,305,532–1,359,224 nt and 4,109,397–4,157,770 nt.), showing 30% and 40% similarity with gene clusters for turnerbactin [47] and the lipopeptide ralsolamycin, respectively. Ralsolamycin has been firstly isolated from *Ralstonia solanacearum* and has been shown to facilitate bacterial entry and induce chlamydospore formation in fungi [48]. In *Kl. pasteurii* BDA134-6, a cluster with 100% similarity with that of the biosynthesis of kleboxymycin [49] (3,081,276–3,127,465 nt), and an additional cluster with 30% similarity with the gene cluster for turnerbactin was found (1,594,136–1,645,420 nt). Very recently, a report on the diffusion of kleboxymycin in different *K. oxytoca* strains was posted [50]. The presence of such a gene cluster in a plant-associated bacterium deserves further attention on a role in endophytic colonization and on gene flow of enterotoxins between environmental and clinical strains of *Klebsiella*.

Table 2. Identification of *nif* gene cluster in the genome of *Kl. pasteurii* BDA134-6, *Ko. pseudosacchari* BDA62-3 and *P. diazotrophicus* BDA59-3.

Gene Code	Gene Name
Kl. pasteurii BDA134-6	
BDA134-6_02421	*nifJ*
BDA134-6_03551	*nifH*
BDA134-6_03552	*nifD*
BDA134-6_03553	*nifK_1*
BDA134-6_03556	*nifD_2*
BDA134-6_03557	*nifK_2*
BDA134-6_03560	*nifS*
BDA134-6_03562	*nifW*
BDA134-6_03565	*nifF*
BDA134-6_03566	*nifL*

Table 2. *Cont.*

Gene Code	Gene Name
BDA134-6_03567	*nifA*
BDA134-6_03568	*nifB*
Ko. pseudosacchari BDA62-3	
Entero_BDA62.3_02197	*nifQ*
Entero_BDA62.3_02198	*nifB*
Entero_BDA62.3_02199	*nifA*
Entero_BDA62.3_02200	*nifL*
Entero_BDA62.3_02201	*nifF*
Entero_BDA62.3_02203	*nifZ*
Entero_BDA62.3_02204	*nifW*
Entero_BDA62.3_02206	*nifS*
Entero_BDA62.3_02207	*nifU*
Entero_BDA62.3_02208	Dinitrogenase iron-molybdenum cofactor
Entero_BDA62.3_02209	*nifK_1*
Entero_BDA62.3_02210	*nifD_1*
Entero_BDA62.3_02211	Dinitrogenase iron-molybdenum cofactor
Entero_BDA62.3_02212	*nifT/fixU*
Entero_BDA62.3_02213	*niK_2*
Entero_BDA62.3_02214	*nifD_2*
Entero_BDA62.3_02215	*nifH*
P. diazotrophicus BDA59-3	
Citro_BDA59-3_02396	*nifJ_1*
Citro_BDA59-3_02623	*nifJ_2*
Citro_BDA59-3_02624	*nifH*
Citro_BDA59-3_02625	*nifD_1*
Citro_BDA59-3_02626	*nifK_1*
Citro_BDA59-3_02627	*nifT/fixU*
Citro_BDA59-3_02628	*Dinitrogenase iron-molybdenum cofactor*
Citro_BDA59-3_02629	*nifD_2*
Citro_BDA59-3_02630	*nifK_2*
Citro_BDA59-3_02631	*Dinitrogenase iron-molybdenum cofactor*
Citro_BDA59-3_02632	*nifU*
Citro_BDA59-3_02633	*nifS*
Citro_BDA59-3_02635	*nifW*
Citro_BDA59-3_02636	*nifZ*
Citro_BDA59-3_02638	*nifF*
Citro_BDA59-3_02639	*nifL*
Citro_BDA59-3_02640	*nifA*
Citro_BDA59-3_02641	*nifB*

3.4. N-Fixing Endophytes Differently Colonize the Primary (O. glaberrima) and Secondary (O. sativa) Host

The ability of the N-fixing endophytes to individually colonize *O. glaberrima* RAM133 and *O. sativa* cv. Baldo plants have been evaluated by measuring the colony-forming units (CFU) of homogenates from whole inoculated plants. The data reported in Table 3 showed that the N-fixing strains colonize more efficiently *O. glaberrima* plants than the *O. sativa* ones. Different colonization ability of *O. sativa* plants was observed for the selected diazotrophic strains: the highest CFU number was obtained for the strains *Kl. pasteurii* 134-6, *E. sacchari* BDA86-11 and *Ko. pseudosacchari* BDA62-3, while the one estimated for the strain, *P. diazotrophicus* BDA59-3 was about a third lower. The lowest colonization ability was observed for the strains *Ko. oryzendophytica* BDA137-1 and *Enterobacter* sp. BDAM41-2. The strains *Kl. pasteurii* BDA134-6, *Ko. pseudosacchari* BDA62-3, and *P. diazotrophicus* BDA59-3 were selected for further analysis.

Table 3. Endophytic colonization of *O. glaberrima* and *O. sativa* plants by six N-fixing endophytes.

Strain	*O. glaberrima* (CFU per Plant) § ($\times 10^5$)	*O. sativa* (CFU per Plant) § ($\times 10^5$)
Ko. oryzendophytica BDA137-1	151 ± 24 ^{e}	10 ± 2 ^{d}
Enterobacter sp. *BDAM41-2*	2264 ± 410 ^{a}	10 ± 2 ^{d}
Kl. pasteuri BDA134-6	1064 ± 93 ^{b}	176 ± 39 ^{a}
E. sacchari BDA86-11	613 ± 76 ^{c}	180 ± 38 ^{a}
Ko. pseudosacchari BDA62-3	312 ± 92 ^{d}	154 ± 24 ^{b}
P. diazotrophicus BDA59-3	384 ± 57 ^{d}	49 ± 8 ^{c}

§ The root of germinated seeds of *O. glaberrima* and *O. sativa* were cut (0.5 cm from the bottom) with a sterile bistoury and incubated in Petri dishes with 50 mL of 1× PBS solution containing each strain to a final concentration of 10^6 cells mL^{-1} for 4 h at room temperature. Inoculated seeds were grown in sand perlite soil. At 14 days after the inoculation, the endophytes were re-isolated from the root tissues to evaluate the number of CFU per plant. The values reported in the table are the mean ± SD of at least five independent replicates each carried out at different times. Means with different letters are significantly different at 5% level of confidence.

3.5. Colonization of Rice Roots by the Strains Ko. pseudosacchari BDA62-3, Kl. pasteurii BDA134-6 and P. diazotrophicus BDA59-3

The use of fluorescently tagged bacteria and CLSM enabled to demonstrate that the strains *Ko. pseudosacchari BDA62-3, Kl. pasteurii BDA134-6* and *P. diazotrophicus BDA59-3* effectively colonize roots of in vitro rice plants under gnotobiotic conditions. The entire colonization process by tagged *Ko. pseudosacchari BDA62-3, Kl. pasteurii BDA134-6* and *P. diazotrophicus BDA59-3* derivatives were recorded, from bacteria adhesion to the root surface to the colonization of the intercellular spaces within the cortex of the differentiation zone (Figure 3). The rice root rhizoplane was effectively colonized by the three strains. The bacteria were predominantly detected on the rice root differentiation.

The observations showed that the cells of all bacteria strains were attached along the whole root surface, mostly at random distribution, displaying no preferred site for the external colonization (Figure 3a–c). Internal colonization of root hairs was also detected (Figure 3a–c). Inner colonization of root hair cells did not occur frequently and only a low proportion was internally colonized. From epidermal cells, bacteria were always detected in the intercellular spaces of the root cortex and vascular tissues, never inside cortical cells or invading the interior of the vascular tissue for the three analysed strains (Figure 3d–f). No significant changes were observed during the bioassay for all the rice plants analysed. It is worth mentioning that the tagged bacteria *Ko. pseudosacchari BDA62-3, Kl. pasteurii BDA134-6* and *P. diazotrophicus BDA59-3* were never detected in the shoots (Figure 3g–o).

3.6. The Endophytes Kl. pasteurii BDA134-6 and P. diazotrophicus BDA59-3 Improve N-Fixation in Two Host Rice Species

To evaluate the N-fixing ability of six endophytic strains when they interact with their host plants (*O. glaberrima* RAM133 and *O. sativa* cv. Baldo) an ARA test was carried out on two-week-old plants inoculated with the different strains. Results reported in Figure 4 show that the inoculation of the two *Oryza* species with the strains *Kl. pasteurii* BDA134-6 and *P. diazotrophicus* BDA59-3 led to the highest nitrogenase activity. The lowest nitrogenase activity was observed when *Enterobacter* sp. BDAM41-2 and *Ko. oryzendophytica* BDA137-1 were used for the inoculation of *O. glaberrima* and *O. sativa* plants, respectively.

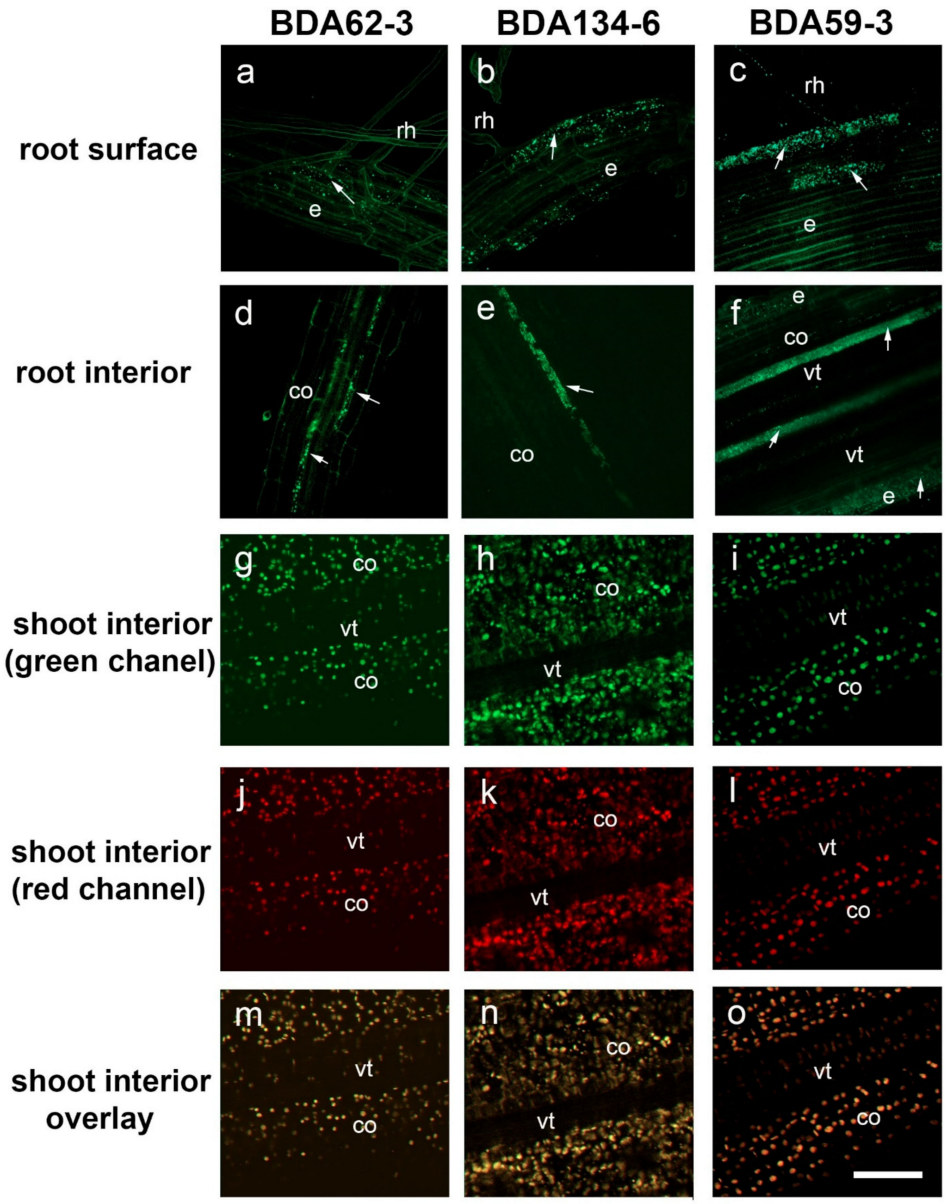


Figure 3. Visualization of the tagged strains *Ko. pseudosacchari* BDA62-3, *Kl. pasteurii* BDA134-6 and *P. diazotrophicus* BDA59-3 in *O. sativa* root and shoot tissues using confocal laser scanning microscopy (CLSM). Confocal analysis was performed on whole representative roots (**a**–**f**) and shoots (**g**–**o**) to show surface and inner colonization of the tagged bacterial strains. Images are projections of 10 optical sections for all panels except for panels (**g**–**o**) where projections were made of 20 optical sections. The focal step size was 4 μm. (**a**–**c**) Detection of BDA62-3, BDA134-6 and BDA59-3 bacteria on the root surface and internal colonization of epidermal cells, including the modified epidermal cells (root hair cells). (**d**–**f**) The internal colonization of the intercellular spaces in the cortex and vascular tissues of the root differentiation zone. The interior of the vascular tissue was never observed invaded by any of the different bacterial strains. (**g**–**o**) Images of the shoots' interior showing the cortex and the vascular tissues. (**g**–**i**) Bacteria were never detected in the shoots but chloroplasts, which are highly autofluorescent, were also detected using both the green (**g**–**o**) and the red channel (**j**–**l**). (**m**–**o**) Overlay of green and red panels of the rice shoots for the three bacteria strains showing overlay of the fluorescent signals for chloroplasts. No bacteria were visualized in green in any case. Bacteria colonization is arrowed. co, cortical cells; e, epidermal cells; rh, root hairs; vt, vascular tissue. Scale bar represents 50 μm.

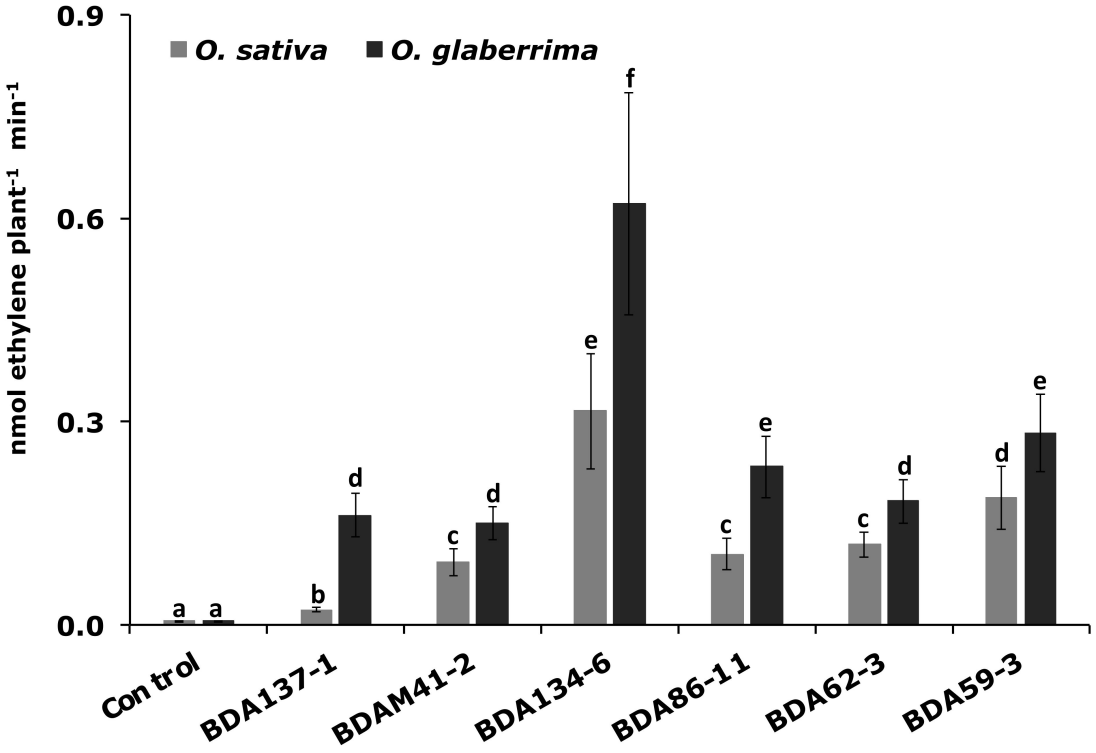


Figure 4. Nitrogenase activity in *O. glaberrima* RAM133 and *O. sativa* L. cv. Baldo plants inoculated with N-fixing endophytes. Two-week-old rice plants of primary (*O. glaberrima*) and secondary (*Oryza sativa*) inoculated host were carefully removed from the pots and used for the ARA test. Data are the means ± SD of at least six independent biological replicates each carried out at different times. Means with different letters are significantly different at 5% level of confidence (Tukey's post hoc test).

3.7. N2 Gas Incorporation

In addition to the acetylene reduction assay, which is an indirect method to verify nitrogen fixation of endophytes in plants, we have also designed a ^{15}N dinitrogen incorporation experiment to follow ^{15}N2 gas into rice plants through the endophytic strains *Ko. pseudosacchari* BDA62-3, *Kl. pasteurii* BDA134-6 and *P. diazotrophicus BDA59-3*. When two-week-old plants were analyzed a significant (as a value greater than the declared measurement uncertainty) ^{15}N incorporation from ^{15}N gas was detected for rice plants inoculated with *P. diazotrophicus BDA59-3* ($\delta^{15}N = 8.9 \pm 0.2$) as compared to uninoculated ones ($\delta^{15}N = 6.9 \pm 0.2$). For plants inoculated with *Ko. pseudosacchari* BDA62-3 ($\delta^{15}N = 7.1 \pm 0.2$) and *Kl. pasteurii* BDA134-6 ($\delta^{15}N = 7.2 \pm 0.6$) the ^{15}N incorporation was not statistically significant. This result indicated that the strain *P. diazotrophicus* BDA59-3 transferred the fixed nitrogen to *O. sativa* plants.

3.8. Kl. pasteurii BDA134-6 and P. diazotrophicus BDA59-3 Increase Carbon (C), Nitrogen (N) and Chlorophyll (Chl) Content of Inoculated O. sativa Plants

To analyse the effect of the inoculation with the strains *E. sacchari* BDA62-3, *Kl. pasteurii* BDA134-6 and *P. diazotrophicus* BDA59-3 on the growth of *O. sativa* L. cv. Baldo plants the total content of carbon (C), nitrogen (N) and chlorophyll (Chl) were measured under hydroponic conditions at 14 and 40 DAI. Uninoculated plants were used as reference (Figure 5). A significant increase in all three parameters was measured for plants inoculated with the strain *P. diazotrophicus.* BDA59-3 at both 14 and 40 DAI. When 'Baldo' plants were inoculated with *Kl. pasteurii* BDA134-6, the C and N content increased only at 40 DAI, and the measured levels were lower than the ones estimated for plants inoculated with *P. diazotrophicus* BDA59-3. Plants inoculated with the strain *E. sacchari* BDA62-3 did not show significant changes in any of the analysed parameters both at 14 and 40 DAI.

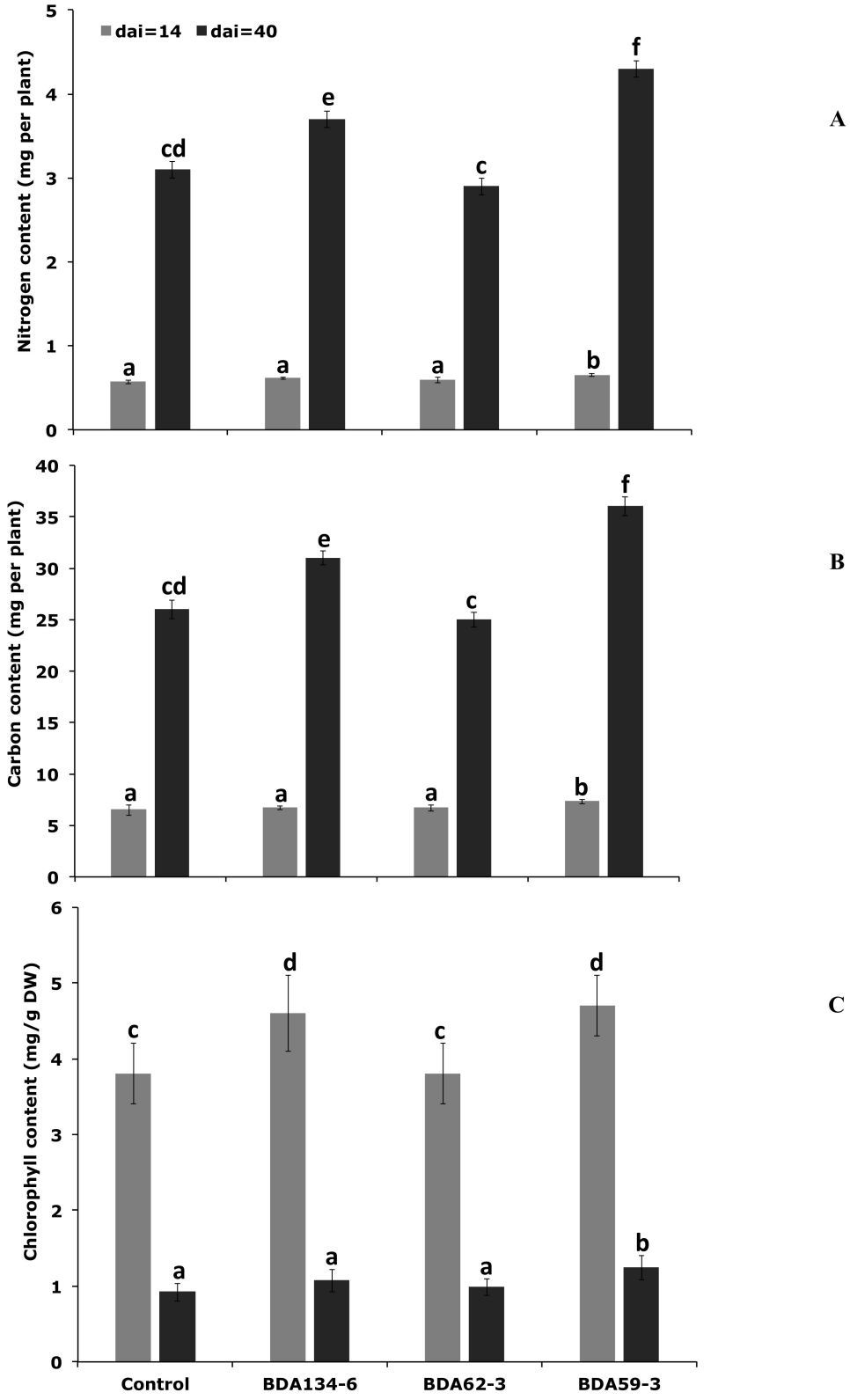


Figure 5. Effect of inoculation with *Ko. pseudosacchari* BDA62-3, *Kl. pasteurii* BDA134-6 and *P. diazotrophicus* BDA59-3 on physiological parameters of *O. sativa* L. cv. Baldo plants. *O. sativa* plants were inoculated with the individual strains and grown in hydroponic conditions as described in Material and Methods section. At 14 and 40 DAI, plants were removed from the growth unit and the leaves detached and used for the analyses of N (**A**), C (**B**), and chlorophyll (**C**) content. Data are the means ± SD of at least five independent biological replicates each carried out at different times. Means with different letters are significantly different at 5% level of confidence (Tukey's post hoc test).

3.9. Inoculation of Rice Plants with Kl. pasteurii BDA134-6 Allows Maintenance of the Nitrogenase Function and Increases the Activity of Antioxidative Enzymes under Salinity Stress

'Baldo' rice was single-inoculated with the endophytes *Kl. pasteurii* BDA134-6 (good IAA-producer and N-fixer) and *P. diazotrophicus* BDA59-3 (weak IAA-producer and N-fixer) and subjected to salinity stress at 21 DAI. The effect of NaCl treatment on N-fixation and stress response was then evaluated in inoculated plants by using the unstressed plants as reference. After salinity stress, plants inoculated with *Kl. pasteurii* BDA134-6 showed only a slight repression of nitrogenase activity (Figure 6). Instead, for *P. diazotrophicus* BDA59-3-inoculated plants, a significant decrease of nitrogenase activity was estimated (Figure 6). Consistently, the activity of all antioxidative enzymes was significantly enhanced in the tissues of plants inoculated with *Kl. pasteurii* BDA134-6 (Figure 7). For *P. diazotrophicus* BDA59-3-inoculated plants, a positive effect was observed only for the activity of APX (Figure 7).

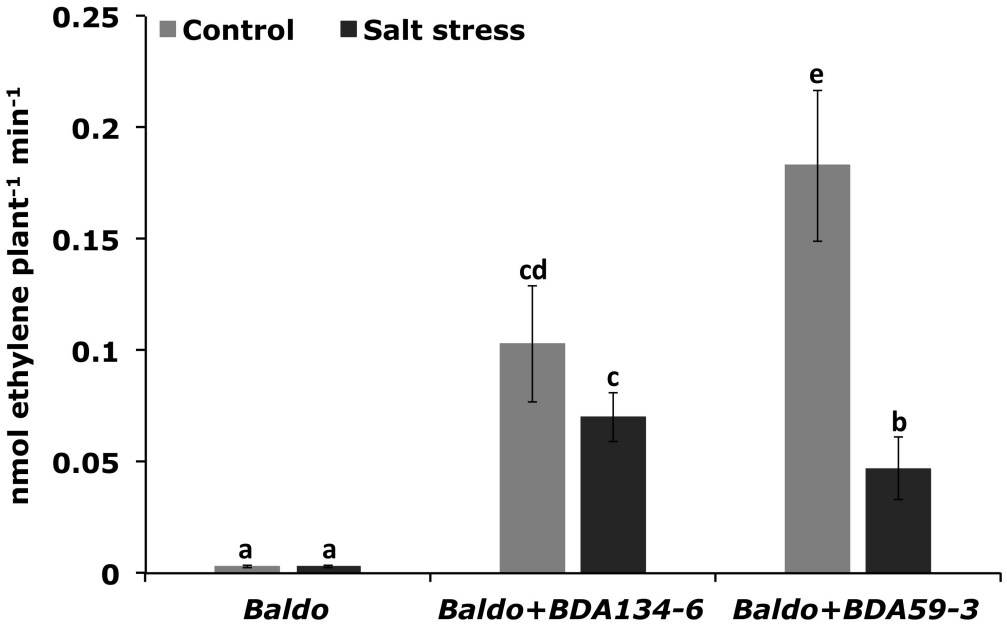


Figure 6. Effect of salinity stress on nitrogenase activity of *O. sativa* L. cv. Baldo plants inoculated with the endophytes *Kl. pasteurii* BDA134-6 and *P. diazotrophicus* BDA59-3. Three-week-old rice plants inoculated with the selected strains were carefully removed from the pots, treated with 0.3 M NaCl for 24 h and used for the ARA test. Data are the means ±SD of at least six independent biological replicates, each carried out at different times. Means with different letters are significantly different at 5% level of confidence (Tukey's post-hoc test).

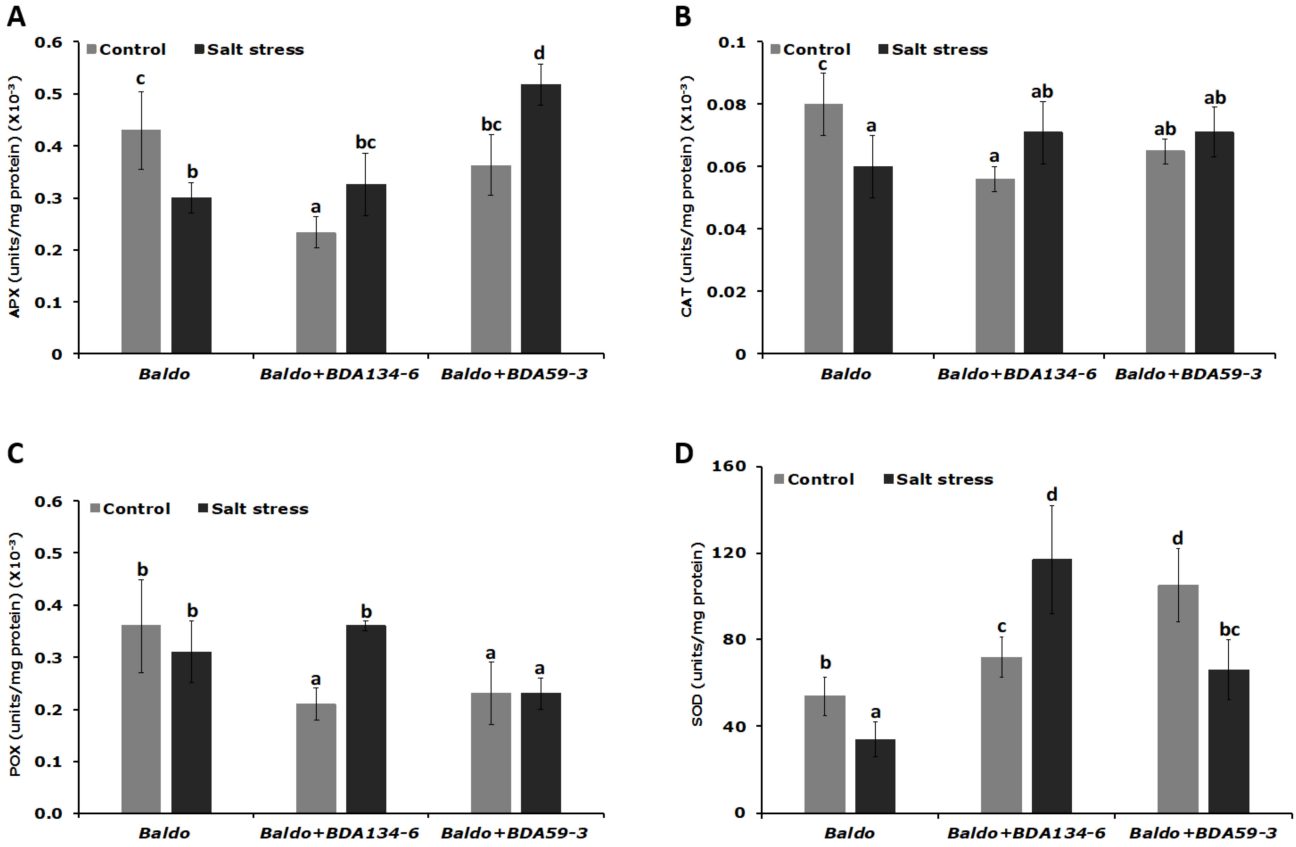


Figure 7. Impact of salinity stress on the activity of the anti-oxidative enzyme in *Kl. pasteurii* BDA134-6- and *P. diazotrophicus* BDA59-3-inoculated plants. Three-week-old rice plants inoculated with the strains *Kl. pasteurii* BDA134-6 and *P. diazotrophicus* BDA59-3 were treated with 0.3 M NaCl for 24 h. After NaCl treatment, the soluble proteins were extracted from leaves and the activity of the enzymes ascorbate peroxidase (APX) (**A**), catalase (CAT) (**B**), peroxidase (POX) (**C**) and superoxide dismutase (SOD) (**D**) measured. Data are the means ± SD of at least five independent biological replicates each carried out at different times. Means with different letters are significantly different at 5% level of confidence (Tukey's post-hoc test).

4. Discussion

Because of their domestication, the human interventions in agriculture and the change in environmental conditions plants might have lost their ability to symbiotically associate with endophytic microbes supporting their response against biotic and abiotic stresses [51,52].

In our study, we tested the ability of the stable crop Asiatic rice *O. sativa* L. to exploit the PGP and stress alleviation properties of *O. glaberrima* bacterial endophytes. For this aim, *O. sativa* cv. 'Baldo' rice, extensively cultivated in Northern Italy (with a range of temperatures 0–16 °C in winter and 5–29 °C in summer), was inoculated with bacterial endophytes isolated from African rice *O. glaberrima* plants cultivated in the Niger River valley (with a range of temperatures 18–27 °C in winter and 30–40 °C in summer). This region is characterized by degraded soil quality and severe environmental stresses, such as heat and extreme variation in the water regime.

O. sativa and *O. glaberrima* are both domesticated rice strains, and *O. glaberrima* has far lower productivity compared to *O. sativa.* However, *O. glaberrima* is cultivated with poor or no agronomic inputs, it has acquired and developed tolerance and resistance against the majority of the biotic and abiotic stresses affecting rice cultivation in Sub-Saharan Africa, and is probably more prone to associate with endophytes promoting plant growth.

The most common endophytic orders detected in this work were the Enterobacterales and Pseudomonadales followed by Xanthomonadales and Bacillales. These results are consistent with other studies concerning the isolation of endophytic bacterial communities

from host plants by culture-based methods [53,54] and with the bacterial diversity pattern previously reported for rice microbiome in other studies [55,56].

It is well known that most of the endophytic bacteria possess features that are directly or indirectly involved in promoting plant growth [57]. For instance, the production of indole-3-acetic acid (IAA) is an important trait through which endophytic bacteria communicate with the host plants enhancing their growth [58]. IAA is the main auxin acting as phytohormone in many plant developmental processes [59]. Auxin homeostasis has a major impact on root architecture, thereby improving nutrient acquisition [60]. A high percentage of endophytic bacteria isolated in this study produced IAA. This result is in agreement with those reported in the literature, reinforcing the concept that the production of IAA is very common among bacteria with endophytic behavior [61,62].

Another important feature of PGPB is the ability to fix nitrogen. Endophytic bacteria can supply fixed atmospheric nitrogen to their host plant by expressing the nitrogenase, a highly conserved enzymatic complex. These bacteria are commonly known as endophytic diazotrophs. Although, endophytic diazotrophs are less efficient in nitrogen-fixation ability than the legume symbiont rhizobia, some N-fixing strains have been identified [58]. In this study, six nitrogen-fixing strains having the *nifH* gene were identified: *Ko. oryzendophytica* (BDA137-1), *E. sacchari* (BDA86-11), *Ko. pseudosacchari* (BDA62-3), *Kl. pasteurii* (BDA134-6), *Enterobacter* sp. (BDAM41-2) and *P. diazotrophicus* (BDA59-3). The presence of the full gene set for nitrogen fixation was confirmed by genome sequencing of three of them (*Ko. pseudosacchari* BDA62-3, *Kl. pasteurii* BDA134-6, and *P. diazotrophicus* BDA59-3). The diazotrophic nature of these endophytes was confirmed by the ARA test performed on free-living strains. The strain *Kl. pasteurii* BDA134-6 had the highest nitrogenase activity. As the nitrogenase enzyme complex in free-living diazotrophs is highly sensitive to molecular oxygen, we speculate that in *Kl. pasteurii* BDA134-6, as well as in the other five isolated diazotrophs, the activity of largely uncoupled respiratory electron transport systems may have protected the nitrogenase enzyme from oxygen damage, as already reported in other studies [63]. Further studies will provide additional data to test such hypothesis.

In the present work, the marker gene *acdS*, which codes for the 1-aminocyclopropane-1-carboxylate (ACC) deaminase, was also identified in six strains belonging to the families of the Ralstoniaceae, Microbacteriaceae, Enterobacteriaceae and Xanthomonadaceae. The presence of *acdS* positive strains in the endosphere of *O. glaberrima* is consistent with its tolerance and resistance against biotic and abiotic stresses. Indeed, lowering ethylene levels by 1-aminocyclopropane-1-carboxylate (ACC) deaminase is considered one of the major mechanisms employed by PGPB to promote plant growth under stress conditions [10]. When the nitrogen-fixing abilities of the six N-fixing strains were evaluated in the host plants *O. glaberrima* RAM133 and *O. sativa* cv. Baldo, the highest nitrogenase activity, was recorded for plants inoculated with the strains *Kl. pasteurii* BDA134-6 and *P. diazotrophicus* BDA59-3, for both rice species. It was also found that these strains were able to colonize *O. sativa* cv. Baldo plants quite efficiently, although the CFU number measured was 10 times lower than the one registered for *O. glaberrima* plants inoculated with the same strains. The higher number of CFU measured for *O. glaberrima* inoculated plants might partially explain the higher ARA efficiency measured for these plants as compared to *O. sativa* inoculated with the same strain.

The different colonization efficiency observed for the two genetically distant rice species (*O. glaberrima* and *O. sativa*) used in this study is not trivial considering that there are contrasting reports about the host specificity of endophytes. Some researchers have reported that endophytes are only able to promote growth of plants closely related to their natural hosts. On the other hand, there are studies concerning the ability of endophytes to promote growth of diverse host plants [58]. Host facthat, as already observed for legume plants tors such as genotype and host phylogenetic relationships play a role in shaping endophytic communities of rice cultivars. Furthermore, it is increasingly evident that environmental conditions such as soil nutrition, moisture, and temperature have a direct influence on endophere communities [64]. *O. glaberrima* and *O. sativa* have adapted to

growth in very different climatic conditions: *O. glaberrima* is adapted to harsh African climate and presents resistance or tolerance to viruses, nematode, bacteria, drought, iron toxicity and high salinity [65,66], while *O. sativa* cv. Baldo plants are grown through a submersion technique in regions of Italy characterized by temperate climate with more than 30 °C of excursion between winter and summer. When the nitrogen-fixing strains *Ko. pseudosacchari* BDA62-3 (weak IAA-producer and N-fixer) and *Kl. pasteurii* BDA134-6 (good IAA-producer and N-fixer) and *P. diazotrophicus* BDA59-3 (weak IAA-producer and N-fixer) were used to inoculate *O. sativa* cv. Baldo we verified that none of them was able to significantly colonize the shoot tissues. By contrast, when the roots were analysed, their colonization efficiency was in the following order: *Kl. pasteurii* BDA134-6 > *Ko. pseudosacchari* BDA62-3 > *P. diazotrophicus* BDA59-3.

Fluorescence microscopy visualization confirmed the ability of *Ko. pseudosacchari* BDA62-3, *Kl. pasteurii* BDA134-6, and *P. diazotrophicus* BDA59-3 to colonize the rice root endosphere. For the strain *Kl. pasteurii* BDA134-6, the internal colonization of root hairs could be linked to the presence of a complete set for the T6SS in the genome of this strain, as the genomic sequence suggested. Indeed, it has been shown that a T6SS mutant of *Kosakonia* sp. KO348 displayed a significant decrease in rice rhizoplane and root endosphere colonization, thus suggesting a role in the host-bacteria colonization or interaction [67]. When the content of C, N and chlorophyll was measured for 'Baldo' plants inoculated with the three selected strains, an increase in all three parameters was observed for *P. diazotrophicus* BDA59-3 both at 14 and 40 DAI. These results, together with the data obtained for the nitrogenase activity and the ^{15}N incorporation, clearly demonstrated that the nitrogen fixed by the strain *P. diazotrophicus* BDA59-3 was directly transferred to *O. sativa* plants from their early stages of growth.

On the other hand, for plants inoculated with *Kl. pasteurii* BDA134-6, the level of N and C content increased only at 40 DAI. We hypothesize that for the strain *Kl. pasteurii* BDA134-6 the transfer of fixed N to the host plant occurs in a later phase of plant growth as compared to *P. diazotrophicus* BDA59-3.

Kinetic studies are required to fully evaluate the accumulation of fixed N in plant tissues over time. The highest nitrogenase activity observed for plants inoculated with *Kl. pasteurii* BDA134-6, which is at the same time a N-fixer and IAA-producer, was consistent with previously published data concerning the positive effect of IAA on nitrogen fixation of rhizobacteria interacting with their host plants [35,38]. Furthermore, the levels of C, N, and Chl measured for plants inoculated with the strain *Ko. pseudosacchari* BDA62-3 were consistent with the low nitrogenase activity and the negligible ^{15}N incorporation measured for this strain. We thus hypothesize that, as already observed for legume plants [68], the nitrogen fixation by endophytes within rice plant tissues is tightly linked to carbon metabolism.

Finally, considering that rice is one of the most salt-sensitive cereals and that the anti-oxidative enzymes scavenge ROS produced as metabolic by-products during osmotic stress and convert them into non-reactive form, the activity of the enzymes APX, CAT, POX and SOD was measured for 'Baldo' plants inoculated with *Kl. pasteurii* BDA134-6 and *P. diazotrophicus* BDA59-3 under salt stress conditions.

Our study shows that the strain *Kl. pasteurii* BDA134-6 performed significantly better in the enhancement of all anti-oxidative enzymes activities compared to strain *P. diazotrophicus* BDA59-3. Consistent with these results, a mild loss of nitrogenase activity was observed for *Kl. pasteurii* BDA134-6-inoculated plants, while a stronger reduction was registered for *P. diazotrophicus* BDA59-3-inoculated ones. We speculate that the positive results observed for the strain *Kl. pasteurii* BDA134-6 could be linked to its ability to produce higher IAA levels. Indeed, the increase in the level of this endogenous molecule could lead, as previously observed for rhizobia strains [39], to alterations in the levels of other important hormones, such as ethylene, which promotes fruit ripening and leaf senescence under abiotic stress. Finally, even if *P. diazotrophicus* BDA59-3 and *K. pasteurii* BDA134-6 belong to species for which pathogenic strains are known [69] and potential toxin genes were

found in their genomes, the positive effects observed for plants inoculated with *Kl. pasteurii* BDA134-6 reinforce the suggestion that the bacterial nanomachine T6SS provides fitness and colonization advantage *in planta* and that its role is not restricted to virulence [70].

5. Conclusions

Our study demonstrated that the culturable endophytes isolated from the African rice *O. glaberrima*, adapted to the harsh environments of West Africa (Mali), were able to colonize quite efficiently the domesticated Asian rice *O. sativa* cv. Baldo, selected to growth in the typical cold and wet environmental conditions of Northern Italy. Our results showed that the inoculation of *O. sativa* cv. Baldo plants with the nitrogen-fixing endophyte *P. diazotrophicus* BDA59-3 positively affected its N-fixation and the content of N, C and chlorophyll. These results support the hypothesis that the efficient endophytic colonizers with good ability to fix N stimulate the C and N metabolism in the host plant in a coordinated manner [71].

Our results also revealed that, under salinity stress, 'Baldo' plants inoculated with the strain *Kl. pasteurii* BDA134-6 showed higher activity of enzymes involved in the stress response. We speculate that this effect could be directly or indirectly related to the production of IAA and to the N-fixing activity observed for *Kl. pasteurii* BDA134-6. Our findings suggest that the recovery of microbial biodiversity from compatible plants grown with minimal human intervention and without the addition of nitrogen fertilizers and its introduction into commercial crops, perhaps as seed treatments, could be an effective strategy to reduce reliance on chemical fertilizers and to enhance the plant's stress response.

For the success of these approaches, it is important to consider that the new inoculated endophytes must balance with the resident microbiome of the domesticated crops, and this might require a few rounds of open field selection. Furthermore, the evaluation of the effects related to the use of any opportunistic pathogens is another important aspect that should be considered before proceeding with open field experiments.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/microorganisms9081714/s1, Figure S1: Position and orientation of the *nif* gene cluster in the genome of *Kl. pasteurii* BDA134-6, *Ko pseudosacchari* BDA62-3 and *P. diazotrophicus* BDA59-3. Order of genes is as from genome coordinates, Table S1: Antibiotic resistance of N-fixing endophytes isolated from *O. glaberrima* plants, Table S2: Identification and characterization of culturable endophytes isolated from *O. glaberrima* plants, Table S3: Genome size and genomic features of N-fixing endophytes isolated from *O. glaberrima* plants, Table S4: Results of Type Strain Genome Server, Table S5: Chemotaxis and flagellar assembly, Folder S1: Genome annotations for N-fixing endophytes isolated from *O. glaberrima* plants, Folder S2: AntiSMASH analysis to look at the presence of gene clusters involved in the biosynthesis of secondary metabolites potentially involved in plant–microbe and microbe–microbe interactions, Dataset S1: KO list and COG0677-COG1609 gene list.

Author Contributions: A.A. carried out the ARA tests, performed the re-isolation of diazotrophs from inoculated plants and contributed to the writing of the manuscript; S.R. carried out the isolation, identification and characterization of endophytes and contributed to the writing of the manuscript; C.F., L.C. and A.M. carried out genome sequencing and bioinformatic analysis; P.P. carried out the confocal microscopy analyses; F.C., O.N. and A.S. grew and selected the *O. glaberrima* plants used in this study; S.P. and M.P. carried out experiments with ^{15}N gas; R.D. and A.M. contributed to the interpretation of the results and to the revision of the manuscript; C.B. designed the research and wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the grant MICRO4Legumes ("Il microbioma vegetale simbionte come strumento per il miglioramento delle leguminose foraggere"), D.M.n.89267 (Italian Ministry of Agriculture). Genome sequencing of strains was partially supported by the grant as "Dipartimento di Eccellenza 2018-2022" awarded to the Dipartimento di Biologia, Università di Firenze, by the Italian Ministry of Education, University and Research (MIUR).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The 16S rRNA gene partial sequences were deposited at NCBI GenBank with the following accession numbers (see also Table S2): MT241164 through to MT241232. Genome sequencing is deposited under Bioproject (PRJNA670042).

Conflicts of Interest: Author Oumar Niangado was employed by the company Syngenta. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Kandel, S.L.; Joubert, P.M.; Doty, S.L. Bacterial Endophyte Colonization and Distribution within Plants. *Microorganisms* **2017**, *5*, 77. [CrossRef]
2. Liu, H.; Carvalhais, L.C.; Crawford, M.; Singh, E.; Dennis, P.G.; Pieterse, C.; Schenk, P.M. Inner Plant Values: Diversity, Colonization and Benefits from Endophytic Bacteria. *Front. Microbiol.* **2017**, *8*, 2552. [CrossRef] [PubMed]
3. Olanrewaju, O.; Glick, B.R.; Babalola, O.O. Mechanisms of action of plant growth promoting bacteria. *World J. Microbiol. Biotechnol.* **2017**, *33*, 197. [CrossRef] [PubMed]
4. Riyazuddin, R.; Verma, R.; Singh, K.; Nisha, N.; Keisham, M.; Bhati, K.K.; Kim, S.T.; Gupta, R. Ethylene: A Master Regulator of Salinity Stress Tolerance in Plants. *Biomolecules* **2020**, *10*, 959. [CrossRef]
5. Gupta, B.; Huang, B. Mechanism of Salinity Tolerance in Plants: Physiological, Biochemical, and Molecular Characterization. *Int. J. Genom.* **2014**, *2014*, 701596. [CrossRef] [PubMed]
6. Van Zelm, E.; Zhang, Y.; Testerink, C. Salt tolerance mechanisms of plants. *Ann Rev. Plant Biol.* **2020**, *71*, 403–433. [CrossRef]
7. Sharma, P.; Jha, A.B.; Dubey, R.S.; Pessarakli, M. Reactive Oxygen Species, Oxidative Damage, and Antioxidative Defense Mechanism in Plants under Stressful Conditions. *J. Bot.* **2012**, *2012*, 217037. [CrossRef]
8. Gupta, S.; Pandey, S. ACC Deaminase Producing Bacteria with Multifarious Plant Growth Promoting Traits Alleviates Salinity Stress in French Bean (*Phaseolus vulgaris*) Plants. *Front. Microbiol.* **2019**, *10*, 1506. [CrossRef]
9. Kruasuwan, W.; Thamchaipenet, A. 1-Aminocyclopropane-1-carboxylate (ACC) Deaminase-Producing Endophytic Diazotrophic *Enterobacter* sp. EN-21 Modulates Salt–Stress Response in Sugarcane. *J. Plant Growth Regul.* **2018**, *37*, 849–858. [CrossRef]
10. Tavares, M.J.; Nascimento, F.X.; Glick, B.R.; Rossi, M.J. The expression of an exogenous ACC deaminase by the endophyte Serratia grimesii BXF1 promotes the early nodulation and growth of common bean. *Lett. Appl. Microbiol.* **2018**, *66*, 252–259. [CrossRef]
11. Linares, O.F. African rice (*Oryza glaberrima*): History and future potential. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16360–16365. [CrossRef]
12. Ndjiondjop, M.N.; Wambugu, P.; Sangare, J.R.; Dro, T.; Kpeki, B.; Gnikoua, K. Oryza glaberrima Steud. In *The Wild Oryza Genomes. Compendium of Plant Genomes*; Mondal, T., Henry, R., Eds.; Springer: Cham, Switzerland, 2018; pp. 105–126. [CrossRef]
13. Van Andel, T. African Rice (*Oryza glaberrima* Steud.): Lost Crop of the Enslaved Africans Discovered in Suriname. *Econ. Bot.* **2010**, *64*, 1–10. [CrossRef]
14. Sie, M.; Sanni, K.; Futakuchi, K.; Manneh, B.; Mande, S.; Vodouhe, R.; Dogbe, S.; Dramé, K.-N.; Ogunbayo, A.; Ndjiondjop, M.-N.; et al. Towards a rational use of African rice (*Oryza glaberrima* Steud.) for breeding in Sub-Saharan Africa. *Genes Genomes Genom.* **2012**, *6*, 1–7.
15. Andreozzi, A.; Prieto, P.; Mercado-Blanco, J.; Monaco, S.; Zampieri, E.; Romano, S.; Valè, G.; Defez, R.; Bianco, C. Efficient colonization of the endophytes *Herbaspirillum huttiense* RCA24 and Enterobacter cloacae RCA25 influences the physiological parameters of *Oryza sativa* L. cv. Baldo rice. *Environ. Microbiol.* **2018**, *21*, 3489–3504. [CrossRef]
16. Edwards, J.; Johnson, C.; Santos-Medellin, C.; Lurie, E.; Podishetty, N.K.; Bhatnagar, S.; Eisen, J.A.; Sundaresan, V. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E911–E920. [CrossRef]
17. Santos-Medellin, C.; Edwards, J.; Liechty, Z.; Nguyen, B.; Sundaresan, V. Drought Stress Results in a Compartment-Specific Restructuring of the Rice Root-Associated Microbiomes. *mBio* **2017**, *8*, e00764-17. [CrossRef] [PubMed]
18. Defez, R.; Andreozzi, A.; Bianco, C. The Overproduction of Indole-3-Acetic Acid (IAA) in Endophytes Upregulates Nitrogen Fixation in Both Bacterial Cultures and Inoculated Rice Plants. *Microb. Ecol.* **2017**, *74*, 441–452. [CrossRef]
19. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [CrossRef]
20. Tamura, K.; Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **1993**, *10*, 512–526. [CrossRef] [PubMed]
21. Ardui, S.; Ameur, A.; Vermeesch, J.R.; Hestand, M.S. Single molecule real-time (SMRT) sequencing comes of age: Applications and utilities for medical diagnostics. *Nucleic Acids Res.* **2018**, *46*, 2159–2168. [CrossRef] [PubMed]
22. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [CrossRef]
23. Cuccuru, G.; Orsini, M.; Pinna, A.; Sbardellati, A.; Soranzo, N.; Travaglione, A.; Uva, P.; Zanetti, G.; Fotia, G. Orione, a web-based framework for NGS analysis in microbiology. *Bioinformatics* **2014**, *30*, 1928–1929. [CrossRef]
24. Lee, I.; Chalita, M.; Ha, S.-M.; Na, S.-I.; Yoon, S.-H.; Chun, J. ContEst16S: An algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 2053–2057. [CrossRef]
25. Meier-Kolthoff, J.P.; Göker, M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* **2019**, *10*, 2182. [CrossRef]

26. Grissa, I.; Vergnaud, G.; Pourcel, C. CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* **2007**, *35*, W52–W57. [CrossRef]
27. Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S.Y.; Medema, M.H.; Weber, T. antiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **2019**, *47*, W81–W87. [CrossRef]
28. Moriya, Y.; Itoh, M.; Okuda, S.; Yoshizawa, A.C.; Kanehisa, M. KAAS: An automatis genome annotation and pathway reconstruction server. *Nucleic Acids Res.* **2007**, *35*, W182–W185. [CrossRef]
29. Hui, X.; Chen, Z.; Lin, M.; Zhang, J.; Hu, Y.; Zeng, Y.; Cheng, X.; Ou-Yang, L.; Sun, M.-A.; White, A.P.; et al. T3SEpp: An Integrated Prediction Pipeline for Bacterial Type III Secreted Effectors. *mSystems* **2020**, *5*, e00288-20. [CrossRef] [PubMed]
30. Salkowski, E. Uber das verhalten der skatolcarbonsaure im organismus. *Z. Physiol. Chem.* **1885**, *9*, 23–33.
31. Gaby, J.; Buckley, D.H. A Comprehensive Evaluation of PCR Primers to Amplify the nifH Gene of Nitrogenase. *PLoS ONE* **2012**, *7*, e42149. [CrossRef] [PubMed]
32. Li, Z.; Chang, S.; Ye, S.; Chen, M.; Lin, L.; Li, Y.; Li, S.; An, Q. Differentiation of 1-aminocyclopropane-1-carboxylate (ACC) deaminase from its homologs is the key for identifying bacteria containing ACC deaminase. *FEMS Microbiol. Ecol.* **2015**, *91*, fiv112. [CrossRef] [PubMed]
33. Gao, S.; Xiao, Y.; Xu, F.; Gao, X.; Cao, S.; Zhang, F.; Wang, G.; Sanders, D.; Chu, C. Cytokinin-dependent regulatory module underlies the maintenance of zinc nutrition in rice. *New Phytol.* **2019**, *224*, 202–215. [CrossRef] [PubMed]
34. Bloemberg, G.V.; Wijfjes, A.H.M.; Lamers, G.E.M.; Stuurman, N.; Lugtenberg, B.J.J. Simultaneous Imaging of *Pseudomonas fluorescens* WCS365 Populations Expressing Three Different Autofluorescent Proteins in the Rhizosphere: New Perspectives for Studying Microbial Communities. *Mol. Plant Microbe Interact.* **2000**, *13*, 1170–1176. [CrossRef]
35. Prieto, P.; Mercado-Blanco, J. Endophytic colonization of olive roots by the biocontrol strain Pseudomonas fluorescens PICF7. *FEMS Microbiol. Ecol.* **2008**, *64*, 297–306. [CrossRef] [PubMed]
36. Defez, R.; Esposito, R.; Angelini, C.; Bianco, C. Overproduction of Indole-3-Acetic Acid in Free-Living Rhizobia Induces Transcriptional Changes Resembling Those Occurring in Nodule Bacteroids. *Mol. Plant Microbe Interact.* **2016**, *29*, 484–495. [CrossRef]
37. Stiefel, E.I. The mechanism of nitrogen fixation. In *Recent Development of Nitrogen Fixation*; Newton, W., Postgate, J.R., Rodriguez-Barrueco, C., Eds.; Academic Press: London, UK, 1977; pp. 69–108.
38. Brand, W.A.; Coplen, T.B. Stable isotope deltas: Tiny, yet robust signatures in nature. *Isot. Environ. Health Stud.* **2012**, *48*, 393–409. [CrossRef] [PubMed]
39. Bianco, C.; Defez, R. *Medicago truncatula* improves salt tolerance when nodulated by an indole-3-acetic acid-overproducing *Sinorhizobium meliloti* strain. *J. Exp. Bot.* **2009**, *60*, 3097–3107. [CrossRef]
40. Liang, Y.; Urano, D.; Liao, K.-L.; Hedrick, T.L.; Gao, Y.; Jones, A.M. A nondestructive method to estimate the chlorophyll content of *Arabidopsis* seedlings. *Plant Methods* **2017**, *13*, 26. [CrossRef]
41. Arnon, D.I. Copper enzymes in isolated chloroplasts. polyphenoloxidase in *Beta vulgaris*. *Plant Physiol.* **1949**, *24*, 1–15. [CrossRef]
42. Zhang, G.X.; Peng, G.X.; Wang, E.T.; Yan, H.; Yuan, Q.H.; Zhang, W.; Lou, X.; Wu, H.; Tan, Z.Y. Diverse endophytic nitrogen-fixing bacteria isolated from wild rice *Oryza rufipogon* and description of *Phytobacter diazotrophicus* gen. nov. sp. nov. *Arch. Microbiol.* **2007**, *189*, 431–439. [CrossRef]
43. Pillonetto, M.; Arend, L.N.; Faoro, H.; D'Espindula, H.; Blom, J.; Smits, T.H.; Mira, M.; Rezzonico, F. Emended description of the genus *Phytobacter*, its type species *Phytobacter diazotrophicus* (Zhang 2008) and description of *Phytobacter ursingii* sp. nov. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 176–184. [CrossRef] [PubMed]
44. Borenshtein, D.; Schauer, D.B. The genus *Citrobacter*. In *The Prokaryotes*; Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.H., Stackebrandt, E., Eds.; Springer: New York, NY, USA, 2006. [CrossRef]
45. Bakker, A.; Smith, D.W. Methylation of GATC sites is required for precise timing between rounds of DNA replication in *Escherichia coli*. *J. Bacteriol.* **1989**, *171*, 5738–5742. [CrossRef]
46. Levy, A.; Gonzalez, I.S.; Mittelviefhaus, M.; Clingenpeel, S.; Paredes, S.H.; Miao, J.; Wang, K.; Devescovi, G.; Stillman, K.; Monteiro, F.; et al. Genomic features of bacterial adaptation to plants. *Nat. Genet.* **2017**, *50*, 138–150. [CrossRef]
47. Han, A.W.; Sandy, M.; Fishman, B.; Trindade-Silva, A.E.; Soares, C.A.G.; Distel, D.L.; Butler, A.; Haygood, M.G. Turnerbactin, a Novel Triscatecholate Siderophore from the Shipworm Endosymbiont *Teredinibacter turnerae* T7901. *PLoS ONE* **2013**, *8*, e76151. [CrossRef]
48. Spraker, J.E.; Sanchez, L.M.; Lowe, T.M.; Dorrestein, P.C.; Keller, N.P. *Ralstonia solanacearum* lipopeptide induces chlmydospore development in fungi and facilitates bacterial entry into fungal tissues. *ISME J.* **2016**, *10*, 1317–2330. [CrossRef]
49. Unterhauser, K.; Pöltl, L.; Schneditz, G.; Kienesberger, S.; Glabonjat, R.; Kitsera, M.; Pletz, J.; Prado, F.J.; Dornisch, E.; Lembacher-Fadum, C.; et al. *Klebsiella oxytoca* enterotoxins tilimycin and tilivalline have distinct host DNA-damaging and microtubule-stabilizing activities. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 3774–3783. [CrossRef]
50. Shibu, P.; McCuaig, F.; McCartney, A.L.; Kujawska, M.; Hall, L.J.; Hoyles, L. Improved molecular characterization of the the *Klebsiella oxytoca* complex reveals the presence of the kleboxymycin biosynthesis gene cluster. *bioRxiv* **2020**. [CrossRef]
51. Pérez-Jaramillo, J.E.; Mendes, R.; Raaijmakers, J.M. Impact of plant domestication on rhizosphere microbiome assembly and functions. *Plant Mol. Biol.* **2015**, *90*, 635–644. [CrossRef]
52. Correa-Galeote, D.; Bedmar, E.J.; Arone, G. Maize Endophytic Bacterial Diversity as Affected by Soil Cultivation History. *Front. Microbiol.* **2018**, *9*, 484. [CrossRef] [PubMed]

53. Etminani, F.; Harighi, B. Isolation and Identification of Endophytic Bacteria with Plant Growth Promoting Activity and Biocontrol Potential from wild Pistachio Trees. *Plant Pathol. J.* **2018**, *34*, 208–217. [CrossRef] [PubMed]
54. Brígido, C.; Singh, S.; Menéndez, E.; Tavares, M.J.; Glick, B.R.; Félix, M.D.R.; Oliveira, S.; Carvalho, M. Diversity and Functionality of Culturable Endophytic Bacterial Communities in Chickpea Plants. *Plants* **2019**, *8*, 42. [CrossRef] [PubMed]
55. Bertani, I.; Abbruscato, P.; Piffanelli, P.; Subramoni, S.; Venturi, V. Rice bacterial endophytes: Isolation of a collection, identification of beneficial strains and microbiome analysis. *Environ. Microbiol. Rep.* **2016**, *8*, 388–398. [CrossRef]
56. Moronta-Barrios, F.; Gionechetti, F.; Pallavicini, A.; Marys, E.; Venturi, V. Bacterial Microbiota of Rice Roots: 16S-Based Taxonomic Profiling of Endophytic and Rhizospheric Diversity, Endophytes Isolation and Simplified Endophytic Community. *Microorganisms* **2018**, *6*, 14. [CrossRef] [PubMed]
57. Greetatorn, T.; Hashimoto, S.; Sarapat, S.; Tittabutr, P.; Boonkerd, N.; Uchiumi, T.; Teaumroong, N. Empowering rice seedling growth by endophytic *Bradyrhizobium* sp. SUTN 9-2. *Lett. Appl. Microbiol.* **2019**, *68*, 258–266. [CrossRef]
58. Afzal, I.; Shinwari, Z.K.; Sikandar, S.; Shahzad, S. Plant beneficial endophytic bacteria: Mechanisms, diversity, host range and genetic determinants. *Microbiol. Res.* **2019**, *221*, 36–49. [CrossRef]
59. Ljung, K. Auxin metabolism and homeostasis during plant development. *Development* **2013**, *140*, 943–950. [CrossRef]
60. Olatunji, D.; Geelen, D.; Verstraeten, I. Control of Endogenous Auxin Levels in Plant Root Development. *Int. J. Mol. Sci.* **2017**, *18*, 2587. [CrossRef]
61. Spaepen, S.; Vanderleyden, J.; Remans, R. Indole-3-acetic acid in microbial and microorganism-plant signaling. *FEMS Microbiol. Rev.* **2007**, *31*, 425–448. [CrossRef]
62. Egamberdieva, D.; Wirth, S.; Alqarawi, A.; Abd Allah, E.; Hashem, A. Phytohormones and Beneficial Microbes: Essential Components for Plants to Balance Stress and Fitness. *Front. Microbiol.* **2017**, *8*, 2104. [CrossRef] [PubMed]
63. Hill, S. Physiology of nitrogen fixation in free-living heterotrophs. In *Biological Nitrogen Fixation*; Stacey, G., Burris, R.H., Evans, H.J., Eds.; Chapman Hall: New York, NY, USA, 1992; pp. 87–134.
64. Singha, K.M.; Singh, B.; Pandey, P. Host specific endophytic microbiome diversity and associated functions in three varieties of scented black rice are dependent on growth stage. *Sci. Rep.* **2021**, *11*, 12259. [CrossRef] [PubMed]
65. Ndjiondjop, M.N.; Cisse, F.; Girma, G.; Sow, M.; Bocco, R.; Djedatin, G.; Blandine, F. Morpho-agronomic and molecular characterisation of *Oryza glaberrima* germplasm from Mali. *Afr. J. Biotechnol.* **2010**, *9*, 7409–7417. [CrossRef]
66. Wambugu, P.W.; Ndjiondjop, M.-N.; Henry, R. Advances in Molecular Genetics and Genomics of African Rice (*Oryza glaberrima* Steud). *Plants* **2019**, *8*, 376. [CrossRef] [PubMed]
67. Mosquito, S.; Bertani, I.; Licastro, D.; Compant, S.; Myers, M.P.; Hinarejos, E.; Levy, A.; Venturi, V. In Planta Colonization and Role of T6SS in Two Rice *Kosakonia* Endophytes. *Mol. Plant Microbe Interact.* **2020**, *33*, 349–363. [CrossRef] [PubMed]
68. Udwardi, M.; Poole, P.S. Transport and metabolism in legume-rhizobia symbioses. *Ann. Rev. Plant Biol.* **2013**, *64*, 781–805. [CrossRef]
69. Merla, C.; Rodrigues, C.; Passet, V.; Corbella, M.; Thorpe, H.A.; Kallonen, T.V.S.; Zong, Z.; Marone, P.; Bandi, C.; Sassera, D.; et al. Description of *Klebsiella spallanzanii* sp. nov. and of *Klebsiella pasteurii* sp. nov. *Front. Microbiol.* **2019**, *10*, 2360. [CrossRef]
70. Bernal, P.; Llamas, M.A.; Filloux, A. Type VI secretion systems in plant-associated bacteria. *Environ. Microbiol.* **2017**, *20*, 1–15. [CrossRef]
71. Defez, R.; Andreozzi, A.; Romano, S.; Pocsfalvi, G.; Fiume, I.; Esposito, R.; Angelini, C.; Bianco, C. Bacterial IAA-Delivery into Medicago Root Nodules Triggers a Balanced Stimulation of C and N Metabolism Leading to a Biomass Increase. *Microorganisms* **2019**, *7*, 403. [CrossRef]

Manuscript File Click here to view linked References

Monitoring maize sustainability and quality using co-applied SynComs and biochar: evidence from two-year field experiments

Authors:
M. Gullì[1,2§], L. Cangioli[3§], S. Frusciante[4§], S. Graziano[1], M. Caldara[2], A. Fiore[4], A.M. Klonowski[5], E. Maestri[1,2], A. Brunori[4], A. Mengoni[2], A. Pihlanto[6], G. Diretto[4], N. Marmiroli[1,2*#] and A. Bevivino[4*#]

§These authors equally contributed to the work
#Co-last authors
*Authors to whom correspondence should be addressed:
nelson.marmiroli@unipr.it
annamaria.bevivino@enea.it

Affiliations:

[1]Interdepartmental Center SITEIA.PARMA, University of Parma, Parco Area delle Scienze, 43124 Parma, Italy;
[2]Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parco Area delle Scienze 11/A, 43124 Parma, Italy
[3]Department of Biology, University of Florence, 50019 Sesto Fiorentino, Italy
[4]Department for Sustainability, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, ENEA Casaccia Research Center, Via Anguillarese 301, 00123 Rome, Italy
[5]Exploration & Utilisation of Genetic Resources, Matís ohf., Icelandic Food and Biotech R&D, Vínlandsleið 12, 113 Reykjavík
[6]Natural Resources Institute Finland (Luke), Myllytie 1, 31600 Jokioinen, Helsinki, Finland

Abstract

Adoption of sustainable practices in maize cropping is urgent. Synthetic microbial communities (SynComs) made of Plant Growth Promoting Microorganisms (PGPMs), coupled with biochar from residual biomasses, offer an environmentally compatible alternative to the use of inorganic fertilizers and increase soil fertility. This paper, based on a one-shot pot experiment in controlled conditions, extends in a two-year field trial the monitoring of plant physiology, soil biology and chemistry, and product metabolomics. We report here the synergistic effect of biochar, SynComs, and arbuscular mycorrhizal fungi co-application on soil microbiome, maize growth, and kernel metabolomic. SynComs application did not affect the diversity/richness of soil microbial communities, posing therefore a low risk of long-term impacts on the soil microbial ecology, whereas the physiology of

plants was characterized by higher chlorophyll content, higher cob weight and higher kernel weight. The combination of SynComs with biochar provided an effect on kernel metabolism with enrichment in antistress and health-beneficial metabolites. Moreover, the economic feasibility and environmental impact of the transition towards the adoption of more sustainable biofertilizers versus conventional agriculture products are discussed.

Keywords:

Biochar; Synthetic microbial consortia; Sustainable agriculture; From field to fork

1. Introduction

As the world's population increases, food security becomes an issue and a future challenge. More than 600 million people are expected to be affected by hunger in 2030, also due to the consequences of the COVID-19 pandemic on global food security (FAO et al., 2021); in other words, it is necessary to define and adopt cropping alternatives focused both on yield and sustainability. More sustainable agroecosystems, capable of producing healthy and nutritious food, and to maintain the legacy between generations (Tilman et al., 2002), are pursued. These agroecosystems shall account for resilience to climate changes, reversal of biodiversity loss, and of a substantial reduction of environmental pollutants, and more nutritious and healthy products. The massive use of chemical fertilisers to sustain production was a consequence of the Green Revolution endorsed to ensure food security but overexploiting soils with nutrients and elements that affect soil geochemical cycles and microbial biodiversity (Rockström et al., 2009; Steffen et al., 2015). Chemical fertilizers and pesticides have contributed to increasing agricultural production (Tilman et al., 2002) but have brought along a negative impact on human and environmental health affecting the equilibrium between man, plants, animals, and nature (Bindraban et al., 2015; Chandini et al., 2019; Destoumieux-Garzón et al., 2018; Rong et al., 2021). Therefore, it is necessary to nourish the soil with natural solutions at minimal environmental costs. The application in agriculture of several soil microorganisms, including species of bacteria and fungi (i.e. rhizobia, *Azotobacter* sp., *Bacillus* sp., *Azospirillum* sp., *Aerobacter* sp., *Burkholderia* sp., *Pseudomonas* sp. *Aspergillus* sp., *Penicillium* sp.,

Trichoderma sp., *Glomus* sp.) , can reduce the need for chemical inputs, improve the quality and the safety of harvested products and processed foodstuffs while implementing the provision of beneficial ecosystem services (Banerjee & van der Heijden, 2022; Chaudhary et al., 2022).

Plants thrive and develop in close association with microbial communities, bacteria, archaea, fungi, and protists (Compant et al., 2019) collectively termed "plant microbiota" which in turn play a pivotal role in promoting plant health and growth, contributing to its immune response, phytohormone production, mineral nutrition, and metabolic modulation (Sánchez-Cañizares et al., 2017). Microbial inoculants have been used in agriculture either to complement or replace fertilizers (Mącik et al., 2020; Maitra et al., 2021). Most bioinoculants are made of a single beneficial microbe defined as a plant growth-promoting microorganism (PGPM) or arbuscular mycorrhizal fungi (AMF) (De Palma et al., 2022; de Souza et al., 2020) but they are seldom effective following open-field application, and that is why more field trials are needed.

Efficacy may be affected by variable environmental conditions, coexistence with other 'indigenous' microorganisms, and interaction with the host plant (Trivedi et al., 2020), but also by their viability and persistence in the soil after application. The combination of different microbial genera within the core plant microbiota (Mapelli et al., 2023; Riva et al., 2022) with different functional attributes, also called Synthetic microbial Communities (SynComs), represents a new and powerful strategy (Callens et al. 2022; Vaccaro et al. 2022).

Under unfavourable environmental conditions, SynComs can enhance the function of the soil/plant ecosystem improving crop yield and quality (De Palma et al., 2022; Shayanthan et al., 2022). A cross-talking between PGPMs and plant secondary metabolism was reported: plants can influence their microbiome by secreting metabolites that in turn may impact the metabolome of the host plant (Pang et al., 2021), thus affecting plant gene expression (Graziano et al., 2022).

Biochar is gaining importance as a carbon (C) storage material that can improve soil quality (Beesley et al., 2013), as a soil amendment, and as the carrier of beneficial nutrients and microorganisms (Marmiroli et al., 2018; Zhang et al., 2018). Recently, the application of biochar in pot experiments

was reported as beneficial to maize root growth (Graziano et al., 2022; Yan et al., 2022) by changing some soil physicochemical properties (Sohi et al., 2010) and increasing soil fertility (Ding et al., 2016). The combination of biochar and SynComs is considered an enabling, and effective, biotechnology for green and sustainable agriculture. Many authors underline the requirements of robust experiments in view of the validation of microbial consortia before large-scale applications in sustainable crop production. The research here reported was realized for this purpose since the SynComs utilized were developed in the framework of the H2020 SIMBA project (Grant agreement No. 818431); they were carefully designed (Tabacchioni et al, 2021) and needed to be tested on different crops in different environmental conditions (Graziano et al, 2022; Hett et al, 2023; Riboni et al, 2023; Caldara et al, 2024) to prove their efficacy. SynComs were applied to improve crop yield and quality for a more sustainable maize cropping. The coapplication of biochar, SynComs, and AMF as green biofertilizer, was monitored in field trials performed through two running cropping seasons. Results indicated variable effects of biochar on the microbial composition of the rhizosphere microbiota; in contrast, the plant morphological and production traits indicated a main effect linked to seasonality, although within each growing season a positive effect of biochar, when associated with SynComs, was observed for chlorophyll content and yield traits. The metabolomic analyses here reported were focused on the elaboration of the metabolomic signature determined by the combined application of SynComs, biochar, and arbuscular mycorrhizal fungi (AMF) on maize grown in the field in two consecutive years. The results suggest that the coapplication of biochar, SynComs, and AMF as green biofertilizers can contribute to the development of more sustainable agronomic practices to improve crop yield and quality. The economic and environmental impact of a transition to this new type of fertilizers in agriculture are discussed.

2. Literature review

2.1 Alternative fertilizers in sustainable agriculture: from pot experiments to field trials

Sustainable agriculture is a multifaceted concept that includes three dimensions: economic, environmental, and social. Many approaches can be pursued to develop each dimension, but few are the solutions that can include all of them. Certainly, the adoption of alternative biofertilizers is one of these virtuous strategies which is recognised as one of the future alternatives to chemical fertilisers (Ammar et al., 2023). Among biofertilizers, PGPMs have demonstrated to be effective in improving plant performance in many species (Reed & Glick, 2023) due to their peculiar biological activities. Many of the papers report experiments in controlled conditions with single bacterial or fungi species. This approach has helped in showing the specific bioactivity of single microbial strains which exert the growth promoting activities such as nitrogen fixation, solubilization and uptake of minerals, synthesis of hormones and the indirect effects on pathogenic strains as the synthesis of antibiotics, cell-wall degradation enzymes, iron sequestration (Pavlicevic et al., 2023; Reed & Glick, 2023).

The combination of strains with different bioactivities allows to maximize the beneficial growth promoting effects and the control of adverse effects of biotic and abiotic stresses as reported in many crops (Chaudhary et al., 2022; Pavlicevic et al., 2022).

Biochar derived from plant biomasses has been recognised since several decades as an eco-friendly amendment that can improve soil quality, increase carbon sequestration, and reduce GHG emissions (Murtaza et al., 2023; Vijay et al., 2021) Biochar is considered also a hub facilitating PGPM growth and survival due to its morphology and physico-chemical characteristics (Marmiroli et al., 2022; Tan et al., 2022)

2.2 Pot experiments and field trials (similar or prodromic?)

Characterization of PGPM properties and ability to interact with plant tissues usually requires testing at laboratory level before being transferred to the field. The pot experiment under controlled conditions is the most common procedure, in which it is possible to show the direct effect of individual microbial strains or consortia, without any "environmental" disturbance. A further advantage of pot experiments is the possibility to test in limited space many conditions, in terms of

mode of application, combination with different soil types or fertilization conditions, to characterize plant-PGPM crosstalk fully. Usually, pot experiments last a few weeks, and give results regarding the germination and the vegetative growth stage, which certainly are important stages of the life cycle that may determine the plant's final performance. Some studies have shown that there is low reproducibility between pot experiments and field trials using the same treatments (Nkebiwe et al., 2016; Symanczik et al., 2023); the effect on biomass and yield increase are more evident in pot experiments than in field trials (Herrmann et al., 2022; Schmidt & Gaudin, 2018). There are several reasons that could explain this inconsistency. The most significant factors include the timing and method of application, the physical and chemical properties of the soil, the method of fertilization, and the interaction with the microbial populations in the soil. These interactions could determine biological competition, such as niche overlap, resource availability, and predation. Additionally, biotic and abiotic stresses that are typical in the field environment may also play a role (Symanczik et al., 2023).

2..3 Impact of SynComs and biochar on rhizosphere microbiome in crops

The application of microbial consortia consisting of PGPM, both bacteria and fungi, has numerous advantages over the use of single strains. Individual species may have reduced effectiveness due to the potential competition with the resident species, therefore their application on a broad scale could be restricted (Liu et al., 2023). The effect of microbial consortia in agricultural soil can vary according to the crop, soil chemical composition, environmental conditions, and the soil indigenous microbial community. The main concern is that the application of SynComs could lead to a disturbance of the native microbiome, activating complex interactions between inoculated and native species, and could unpredictably influence the balance of the soil biotic component (Liu et al., 2023). This is very critical considering that soil microbiome functions are positively correlated to microbial diversity and that the latter is strongly connected to one health (Banerjee & van der Heijden, 2022). Arable lands unfortunately display lower microbial diversity (Banerjee et al., 2024). Therefore, a crucial issue of

using microbial consortia is the evaluation of its impact on the resident microbial communities and on the modification of the biodiversity. Well-established methods based on metagenomic analyses and targeted NGS sequencing of bacterial ad fungi conserved sequences allow to verify the composition of the soil and rhizosphere microbiome (Nwachukwu & Babalola, 2022).

Moreover, the combination of PGPMs with biochar could lead to unexpected outcomes that need to be investigated. Indeed, biochar properties impact different soil health indicators, such as pH, soil organic matter, and soil stability, along with microbial biomass, richness, and diversity (Deshoux et al., 2023; He et al., 2021). The influence of biochar alone, or combined with SynComs, is not clear, but this is a crucial aspect to study to exclude negative effects, or possibly identify positive outcomes.

2.4 Impact of SynComs and biochar on crop metabolomic profile: from field to fork

The sustainability of agricultural ecosystems, besides the environmental issues, should consider food security (OECD & FAO, 2023) in its fourth dimensions. The adoption of alternative fertilisers and amendments may affect the production both at quantitative (yield) and qualitative levels (nutritional quality); the first aspect has been the focus of many studies (Chouhan et al., 2021), while the second one needs to be further implemented. The characterization of the metabolomic profile of raw materials, as cereal flour, or seeds, has been extensively applied in the last years to elucidate the changes determined by the interactions between plants and microbes. Most research has focused on the response to pathogens or abiotic stresses, while little research reports the effects of PGPMs on the plant metabolome in terms of bioactive compounds beneficial to both plant and consumer health (Riboni et al., 2023; Saia et al., 2019)

3. Methods

3.1. Plant materials and experimental field trials

Field experiments were carried out during two growing seasons (2020 and 2021) on a silty clay loam soil at the experimental farm ‘Azienda Stuard’ (http://www.stuard.it/), close to Parma (North, Italy),

(Lat. 44_4802300N; Long. 10_1603000E; 58 m above sea level). The experiments followed relevant institutional, national, and European guidelines and legislation. Commercially available seeds of *Zea mays* cv DKC6587 (Dekalb®, Monsanto Agricoltura Italia SpA, Milan, Italy) were sown at a density of 8 plants per m², in 8.4 m² plots, adopting a distance between seeds in rows of 17.8 cm and 70 cm interrow spacing. Nine treatments in four replicates (2020) and twelve treatments in three replicates (2021), as described in paragraph 3.3, were distributed with a split-plot design. The sowing dates for the two years were 10/05/20 and 24/04/2021, respectively. Nitrogen (N) fertilizer (with urea 46%) was applied once for a total of 170 kg/ha in 2020 and of 85 kg/ha in 2021. This amount corresponds respectively to 100% and 50% of a normal fertilization plan for maize. Phosphates were omitted in the two years. Samples of soil between 0–30 cm depth were collected prior to seedbed preparation and at harvest time during both growing seasons, for the physicochemical analysis as presented in Supplementary Table S1, S2. At the end of each growing season, plants were manually harvested: on 7 October 2020 and on 30 September 2021.

3.2. Weather conditions and irrigation scheme

Maximum and minimum temperatures (°C) and precipitation (mm) were recorded daily by an automatic weather station installed near the experimental field (Supplementary Fig. S1). Irrigation was performed with a drip system which has the advantage, especially in the driest periods, of compensating for water shortages during the most critical phases of maize life cycle, considering the real needs of the plant. Drip tapes were placed every two rows (1.5 m).

3.3. Source of biofertilizers and biochar

Two synthetic microbial consortia (MC_B and MC_C) were designed as previously described (Tabacchioni et al., 2021). MC_B includes *Azotobacter vinelandii* DSM 2289, *Bacillus* spp. BV84, *Bacillus amyloliquefaciens* LMG 9814, *Pseudomonas fluorescens* DR5, *Rahnella aquatilis* BB23/T4d. MC_C is composed of *Azotobacter chroococcum* LS132, *B. amyloliquefaciens* LMG

9814, *Pseudomonas fluorescens* DR54, *Burkholderia ambifaria* MCI 7, and *Rahnella aquatilis* BB23/T4d. Each single strain was grown in a pilot scale bioreactor according to the procedure previously described (Tabacchioni et al., 2021). Formulation of MC_B and MC_C was made by adding all single strains in a 1:1 ratio as previously described (Hett et al., 2023) to reach at least $>10^6$ CFU for each strain included in the consortia.

The mycorrhizae AMF (MycAgro lab., Bretenière, France) (http://www.mycagrolab.com) consists of a granular inoculum, containing mineral solid particles (clay, zeolite) and *Rhizophagus intraradices* (10 propagules/g containing spores, mycelium and mycorrhizal root pieces) (Fracasso et al., 2020), that was supplied at 3 g to each plant (0,24 tons/ha).

MICOSAT F (F1) in granular form was provided by CCS Aosta Srl (Aosta, Itlay) (https://www.micosat.it/prodotto/micosat-uno/); it was supplied at 3 g to each plant (0.24 tons/ha).

Biochar (CHAR) produced by pyrolysis (or pyrogasification) of biomasses in oxygen-limited conditions was previously characterized (Marmiroli et al., 2018). It was applied before sowing at a rate of 0.2 kg/m^2 (2 tons/ha) and buried to a depth of 10 cm.

For each microbial treatment, the required number of seeds and the specific product were distributed using a pneumatic seeder machine containing two compartments: one for the seeds and one for the consortia. Each compartment has a nozzle whose opening can be regulated to allow the deposition of one seed at a time and the desired quantity of the granular or powdered material.

The microbial treatments and biochar as carrier were utilised in different combinations, as follows: season 2000: Conventional condition (Control), Micosat (F1), AMF, CHAR, CHAR_AMF, CHAR_F1, CHAR_MC_B_AMF, CHAR_MC_C_AMF, CHAR_C_F1; season 2021: Conventional condition (Control), CHAR, MC_B, MC_C, AMF, MC_B _AMF, MC_C_AMF, CHAR_MC_B, CHAR_MC_C, CHAR_AMF, CHAR_MC_B_AMF, CHAR_MC_C_AMF.

The treatments that were maintained throughout the two seasons were: conventional condition (Control), CHAR, AMF, CHAR_AMF, CHAR_MC_B_AMF, CHAR_MC_C_AMF.

3.4. Morphological, physiological, and production traits

The following traits were measured at different plant growth stages. Relative chlorophyll content (CHL) and leaf transpiration rate (LTR) were measured at vegetative growth stage (V6), flowering (R1) and maturity (R4); plant height (PH) was measured at R1 and R4. At harvest (R6), number of ears (E), plant weight (gr) (PW), stock weight (gr) (SW), ear weight (gr) (EW), production at 15% humidity (t/ha) (Y15), yield per plant (YPL), grain moisture % (GM), test weight (kg/hL) (TW), 1000 grain weight (gr) (TGW) were measured.

Chlorophyll content was measured *in vivo* with the SPAD-502 chlorophyll meter (Konica Minolta Business Solution Italia Spa, Milan, Italy), taking ten measurements along the length of five expanded leaves in each plot. Leaf transpiration rate was measured on five expanded leaves in each plot using the portable AP4 Porometer (Delta-T Devices, Cambridge, UK).

During the entire plant growth cycle, a constant visual inspection was conducted to detect any pathological evidence on the leaves, but no sign of infection was ever detected.

The analysis for the presence of mycotoxins was conducted at harvest on grains evaluating the presence of aflatoxins using the kit B-TeZ ELISA AFLA B1 (BioTeZ, Berlin-Buch Gmbh, Germany). In all samples, the level of aflatoxins was below the limit of detection of the method (data not shown).

Standard statistical analyses of agronomic and physiologic data were performed using the software Past 4.03 (© Copyright Hammer 1999-2020). Data were analyzed by one- way ANOVA and Bonferroni-Holm post hoc test, or by two-way ANOVA and Tukey's post-hoc test. A P value ≤ 0.05 was considered statistically significant.

3.5. Soil analysis

Soil chemical analyses were done according to standardised protocols. For organic C, total N, and phosphorus (P), protocols reported in the "D.M. Politiche Agricole 13/09/99 Met. VII3" were adopted, while for all other elements (calcium, Ca; iron, Fe; zinc. Zn; potassium, K) the protocols

defined by LEI/MP/N. 30 2020 Rev.0 were applied. Soil pH and conducibility (CE) were measured as reported by "D.M. Politiche Agricole 13/09/99 Met. 248".

3.6 Rhizospheric soil DNA extraction and 16S rRNA gene amplicon sequencing

Rhizospheric soils were sampled during the two growing seasons at the end of the trial for the metagenomic analysis (27/07/2020 and 03/09/2020; 15/7/2021 and 01/09/2021). Samples of soil were kept at -80 °C until DNA extraction. DNA was extracted using QIAGEN's new DNeasy PowerSoil Pro Kit (QIAGEN Group, Hilden, Germany) following the manufacturer's instructions. Quantity and quality of DNA were evaluated using the NanoDrop 2000 (Thermo Fisher Scientific, Waltham, MA) and Qubit 4.0 fluorometer (Invitrogen, Carlsbad, CA).

DNA quality was also checked by PCR amplification using 16S rRNA gene primers pairs 314F and 805R that amplify a 500 bp amplicon of V3-V4 region. PCR conditions consisted of 1 min at 94°C followed by 25 cycles of 30 s at 94°C, 30 s at 55°C, 45 s at 72°C with a final incubation for 7 min at 72°C. For each sample, two independent DNA extractions were performed and then pooled in an equimolar ratio, obtaining a composite DNA sample for each plant. Samples were stored at -20° C prior to use. Illumina 16S rRNA gene library construction and sequencing were performed using Illumina MiSeq platform and 300 PairedEnds strategy at Matís ltd. Reykjavik, Iceland. Briefly, amplicons corresponding to the V3-V4 regions of 16S rRNA gene were obtained with the following primers: Pro341F, 5′-CCTACGGGNBGCASCAG-3′, and Pro805R, 5′-GACTACNVGGGTATCTAATCC-3′ (Takahashi et al. 2014). Forward and reverse overhangs (5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-(locus-specific sequence)-3' and 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-(locus-specific sequence)-3',) were added for dual index library preparation. The Illumina MiSeq protocol is available at *https://web.uri.edu/gsc/files/16s-metagenomic-library-prep-guide-15044223-b.pdf*. Raw sequence data reported in this study have been deposited in the National Center for Biotechnology Information (NCBI) "Sequence Read Archive" (SRA), under the project accession number PRJNA1023724.

3.7 Bioinformatic and statistical analysis of 16S rRNA gene sequencing data

The pipeline for 16S rRNA amplicon sequence data analysis followed previously published procedures (Cangioli et al., 2022). Samples with reads lower than 5'000 were discarded. In brief, after trimming, paired-end sequences were clustered into Amplicon Sequence Variants (ASVs) following the DADA2 pipeline (version 1.16) (Callahan et al., 2016). Then sequences were filtered to remove chimeras. Taxonomic assignment was carried out against the SILVA NR99rel138 database (Quast et al., 2013) using 'DECIPHER' R package (version 2.18.1) (Wright et al., 2012) as implementation of DADA2 (SSU version 138 available at: http://www2.decipher.codes/Downloads.html). Annotated ASVs count tables were processed by 'Phyloseq' R package (version 1.34.0) in R environment version 4.0.5 (McMurdie & Holmes, 2012). Rarefaction curve were generated using 'ggplot2' (version 3.3.3) (Wickham, 2011) () and 'ranacapa' (version 0.1.0) (Kandlikar et al., 2018) R packages using the 'ggrare()' function on phyloseq object. Alpha diversity (Shannon and Simpson indices) was calculated and plotted using the function 'diversity()' within 'microbiome' R package (version 1.12.0) (Shetty & Lahti, 2019). Good's coverage and Evenness indices were calculated through the R functions 'goods()' and 'evenness()', respectively, within the 'microbiome' R package (version 1.12.0). Multivariate analysis (Principal Component Analysis) was carried out after a *vst* transformation using the following functions: varianceStabilizingTransformation(), plotPCA(), ASVs. The 'ggplot2' R package (version 3.3.3) was used to generate relative abundance plots. The evaluation of the effect of treatment and the relationships between microbiota taxonomic composition and agronomic and physiological data on maize plants was performed by using permutational multivariate analysis of variance (PERMANOVA). PERMANOVA was carried out using the R packages 'ggplot2' (version 3.3.3), 'vegan' (version 2.5-7) and 'pairwise.Adonis' (version 0.0.1) using the functions 'adonis2()' and 'pairwise.adonis()', respectively. To test the differential ASVs occurrence among microbiota from different treatments, differential counts analysis was performed, using the '*DESeq2'* R package

(version 1.30.0) functions: phyloseq_to_deseq2(), estimateSizeFactors(), DESeq(), results() (Love, Huber, and Anders 2014).

3.8 Metabolomic analysis

Maize kernels were collected at maturity and kept at 4°C before being subjected to global metabolic profiling by high pressure liquid chromatography coupled to high resolution mass spectrometry (HPLC-HRMS) procedures. Maize kernels, besides their amount of carbohydrates (75%), proteins (9-10%), lipids (4-5%), and dietary fiber (7-8%), are also an important source of vitamins (B3, group B, A, E, K) and micronutrients (K, P, magnesium, Mg, Ca, Fe, Zn). Furthermore, whole grain, with germ, endosperm, and bran, contain many phytochemicals, with a unique profile of nutrients when compared with whole grains of other cereals (Siyuan et al., 2018). Semi-polar metabolites were extracted from 20 mg lyophilized, homogeneously ground maize grain tissues with 0.75mL of cold 50% (v/v) methanol, 0.1% (v/v) formic acid, spiked with 5µg/mL formononetin as internal standard. After shaking for 40' at 20Hz using a Mixer Mill 300 (Qiagen, Valencia, C, USA), samples were centrifuged for 15 min at 20,000 g at 4ºC. Then, 0.6mL of supernatants were removed, transferred to HPLC tubes, and subjected to HPLC-HRMS as previously reported (Benincasa et al., 2020). Non-polar metabolites were extracted from 30 mg of lyophilized, homogeneously ground grain tissues using 1mL of 25% (v/v) methanol, 50% (v/v) chloroform, 25% (v/v) 50mm-Tris-HCl, spiked with 10µg/mL DL-α-tocopherol acetate as internal standard. After centrifugation, the organic hypo-phase was dried with a Speed Vac concentrator; the residue was resuspended in ethyl acetate (100µL), transferred to HPLC tubes, and subjected to HPLC-HRMS as previously reported (Benincasa et al., 2020). Subsequently, both untargeted and targeted analyses on mass chromatograms were carried out. Untargeted metabolomics was performed using the SIEVE software (Thermofisher Scientific, Waltham, MA, USA) as previously reported (Dono et al., 2020). Targeted metabolite identification was performed by comparing chromatographic and spectral properties with authentic standards (whenever available) and reference spectra, in house database, literature data, and on the basis of the

m/z accurate masses, as reported in the Pubchem database (http://pubchem.ncbi.nlm.nih.gov/) for monoisotopic mass identification, or on the Metabolomics Fiehn Lab Mass Spectrometry Adduct Calculator (http://fiehnlab.ucdavis.edu/ staff/kind/Metabolomics/MS-Adduct-Calculator/) in the case of adduct detection. Data were analyzed by one- way ANOVA, and Tukey's pairwise t-test. A - *p-value* ≤ 0.05 was considered statistically significant.

4. Results

4.1. Environmental conditions during field trials

Agrometeorological data collected from the field trials in 2020 and 2021 are shown in **Supplementary Fig. S1.** A general trend was observed in 2020 when higher rainfall (especially during pre- and post-anthesis periods), and lower minimum and maximum temperature average values were registered compared to 2021. Plots were drip irrigated as needed to optimise maize production, and, considering both rain and irrigation, the total amounts of water received were similar in both years.

4.2. Soil chemical analysis

The total N content was similar in all plots in both growing seasons **(Supplementary Tables S1, S2**), the total organic C was higher in 2021, but in both years no differences were observed among the treatments. In general, the concentrations of these elements averaged that of a healthy soil. The presence of plants was accompanied by an increase in Electrical Conductivity (EC), dry matter, and the amount of Ca, Fe, Zn and K with respect to bare soil. Overall, the values of Ca measured in 2021 were higher than in 2020. In 2021, the highest Ca values were observed for the treatments CHAR_B_AMF and CHAR_C_AMF, while in 2020, treatment with CHAR was associated with the highest amount of Ca and highest pH value. In 2021, pH values were lower than in 2020 in all conditions. At the concentration of biochar tested, there was no significant increase in pH related to

the fact that soil pH (7.8-8.6) and biochar pH (8-8.2; see biochar A4 in (Marmiroli et al., 2018)) were quite close.

An increase in Ca and a decrease in pH are both providing important agronomic benefits to many soils, at least in the short-medium term (Rawat, Sanwal, & Saxena 2018). Moreover, the addition of biochar to the soil can improve maize growth by stimulating plant N-uptake (Peng et al., 2021). Addition of biochar in low amount did not significantly affect the chemical composition of bulk soil as previously reported (Vassura et al., 2023).

4.3. Agronomic and physiological parameters

Maize developed vigorously throughout the two growing seasons without any sign of phytosanitary problems. Physiological and morphological data from all the plots in both seasons are in **Supplementary Tables S3, S4**; data regarding the six common conditions in the two seasons are shown in **Fig. 1**. The final yield in both seasons was not significantly affected by the treatments or by the reduction in fertilization (50% less N and no P). Significant differences were observed for CHL in 2020. At stage V6, the highest CHL values were observed in plants treated with CHAR_AMF, CHAR_MC_B_AMF and CHAR_MC_C_AMF, as compared to CHAR alone; at stage R4, CHAR determined significantly higher CHL values as compared to Control. No significant differences were observed in 2021, but the treatment CHAR_MC_C_AMF determined the highest values of cob FW, plant yield and TGW (**Supplementary Table S3**).

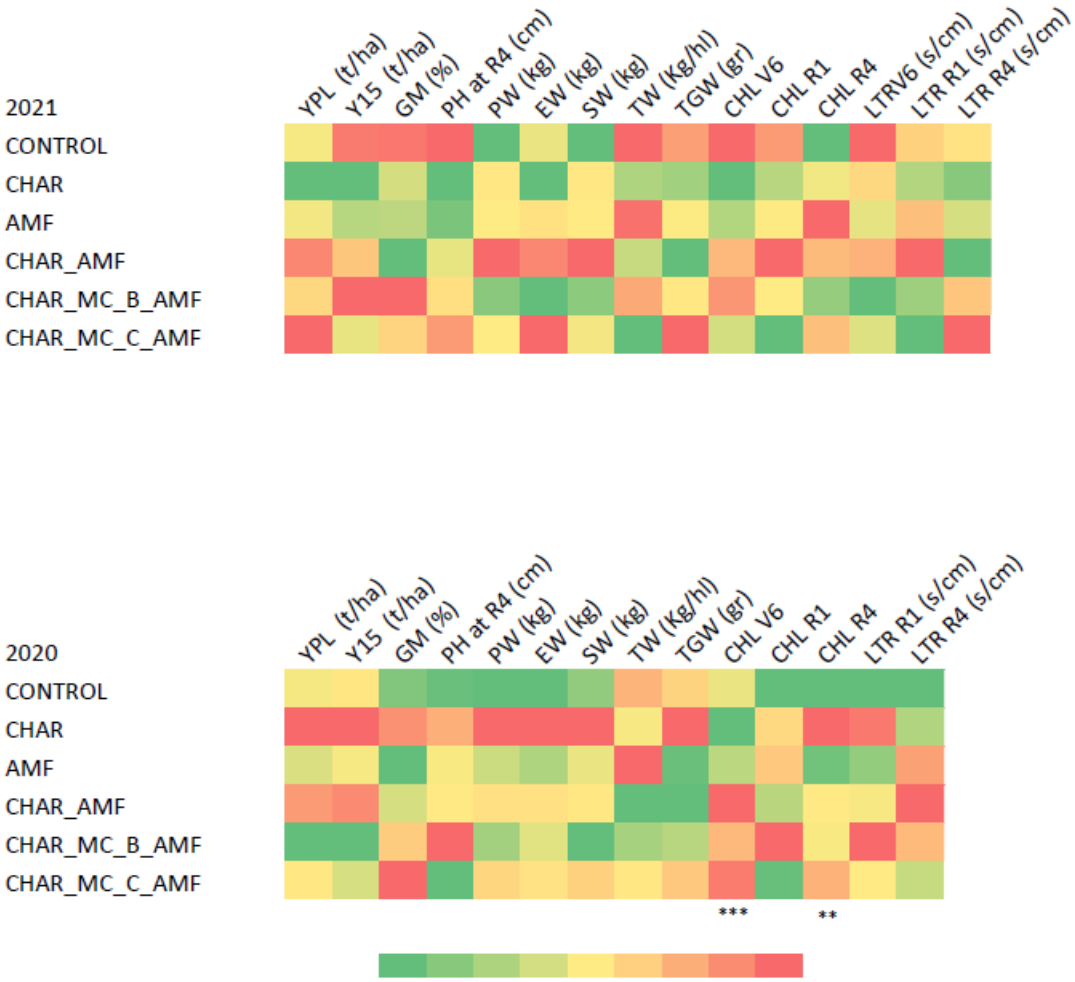


Fig. 1. Agronomic and physiologic data collected from the maize field trials in 2020 (A) and 2021 (B). For each trait, a data standardization was performed, and colour code scale from green to red indicates the variation in the range from -2 to +2. Production per plant (kg) (YPL), Production at 15% humidity (t / ha) (Y15), Grain moisture% (GM), Plant height (cm) (PH), Plant weight (gr) (PW), Ear weight (gr) (EW), Stock weight (gr) (SW), Test weight (Kg/hl) (TW), Thousand grain weight (gr) (TGW). Chlorophyll content (CHL); leaf transpiration rates (s/cm) (LTR). Vegetative growth stage (V6), flowering (R1), maturity (R4). CONTROL- no addition of biofertilizers or biochar; MC_B and MC_C, synthetic microbial consortia; AMF, mycorrhizae; CHAR, Biochar. Asterisks indicate traits showing significant differences among treatments (Anova, Tukey's post hoc test; ** p<0.01, *** p<0.001).

Correlations among plant physiological and agronomic traits and soil composition were studied using Principal Components Analysis (PCA) (**Fig. 2**), this allowed a full comparison between soil chemical composition, plant response, and the formation of the rhizobiome.

PC1 explained 36.9% of the variance and was positively correlated with plant height, Fe and K content in the soil. PC2 explains 25.2% of the variance and was positively correlated with TGW, LTR at maturity, and some soil parameters like Ca, P content and EC. Treatment with

CHAR_MC_C_AMF in both seasons clustered along PC2, separated from the treatments CHAR_AMF and AMF. Treatment with CHAR_MC_B_AMF clustered along PC1 with the control.

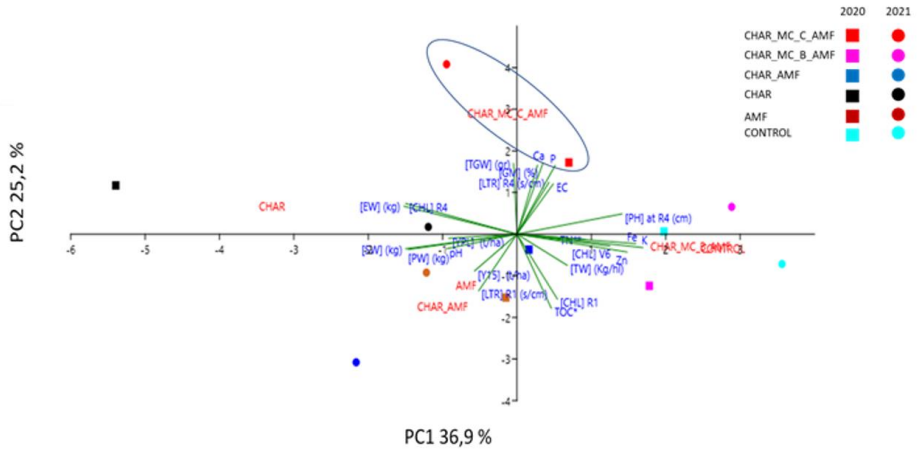


Fig. 2 Principal component analysis (PCA) ordination biplot diagram of maize traits for samples grown under the different treatments in 2020 and 2021. The biplots of the PCA model are calculated on the average values of the investigated parameters for each season. The two first components of the PCA model represented more than half of variation in the dataset (62.1%). The percentage of total variance as explained by each axis is reported. Symbols' legend for treatments are boxed on the top right. CONTROL- no addition of biofertilizers or biochar; MC_B and MC_C, synthetic microbial consortia; AMF, mycorrhizae; CHAR, Biochar.

4.4. Characterization of rhizosphere microbiota

Amplification of 16S rRNA generated 4.383 and 8.482 Amplicon Sequence Variants (ASVs) in 2020 and 2021, respectively (Supplementary Dataset 1). Goods'coverage values of selected samples reached values above 0.99 for all samples, enough to compute reliable alpha-diversity indices. No significant differences in alpha diversity were observed between treatments ($p>0.05$): the application of SynComs did not affect the diversity of the native rhizosphere (Supplementary Table S5).

Concerning the taxonomic distribution of ASVs, those associated with Proteobacteria (18%-20%), Actinobacteria (19%-30%), Chloroflexi (5%-13%), Firmicutes (7%-8%), Verrucomicrobiota (3%-6%), Acidobacteria (7%-8.5%) (Fig. 3A, 3B) were the most abundant in both seasons.

A non-Metric Multidimensional Scaling (nMDS), Permanova and PCA on differently abundant ASVs (DESeq2 analysis) indicated that native microbiota was resilient to treatments, due to the lack

of statistically significant differences among treatments, or times of sampling (Fig. 3C, 3D, Supplementary Fig. S2, Table S6).

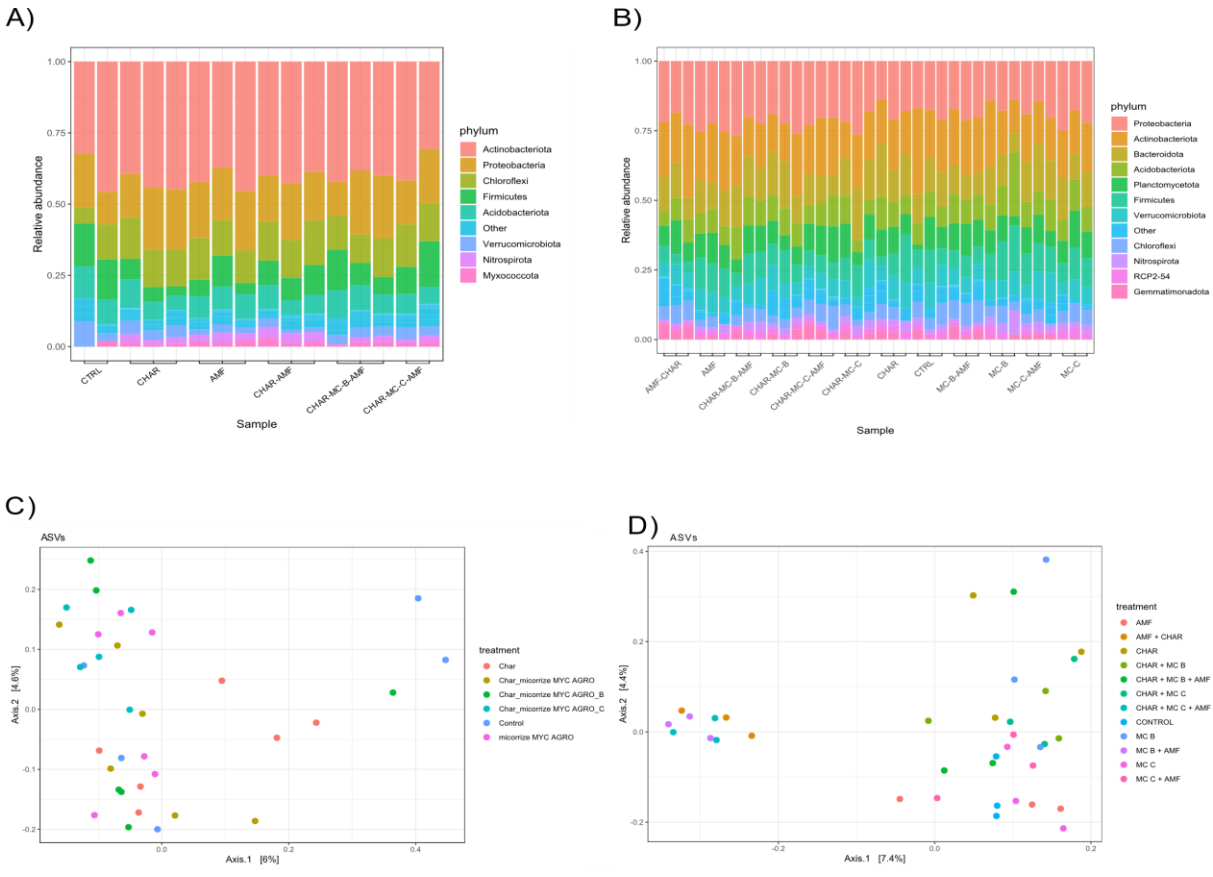


Fig. 3 Relative abundances of taxonomies and rhizospheric maize samples distribution according to bacterial community. Abundances of detected ASVs are reported at phylum level. Panel A) refers to the first season of field trials, 2020, whereas panel B) refers to the second season of field trials, 2021. Principal coordinates analysis (PCoA) of rhizosphere bacterial community was performed using Bray–Curtis distance as ordination method. Colours indicate different treatments applied to maize samples in field trials of season 2020 (C) and 2021 (D). Legends of content of figure A, B, C,D are placed on their right-sides. CONTROL- no addition of biofertilizers or biochar; MC_B and MC_C, synthetic microbial consortia; AMF, mycorrhizae; CHAR, Biochar.

There were no correlations between rhizosphere microbiota and agronomic and physiological data in both seasons at the flowering and harvesting stages (Supplementary Table S6, S7, Fig. S3). Thus, showing only a negligible effect of SynComs over the maize rhizosphere microbiota under the experimental conditions.

4.5. Untargeted metabolomics of maize kernels

The global metabolic profiles of kernels from both seasons were evaluated by HPLC-HRMS. Untargeted analysis of the semi- and non-polar fractions showed putative perturbations in the global metabolome. A set of Differentially Accumulated Metabolites (DAMs) (Supplementary Table S8, S9) was retrieved. PCA gives an overview of the metabolomic performance considering the treatments and the metabolites (Fig. 4).

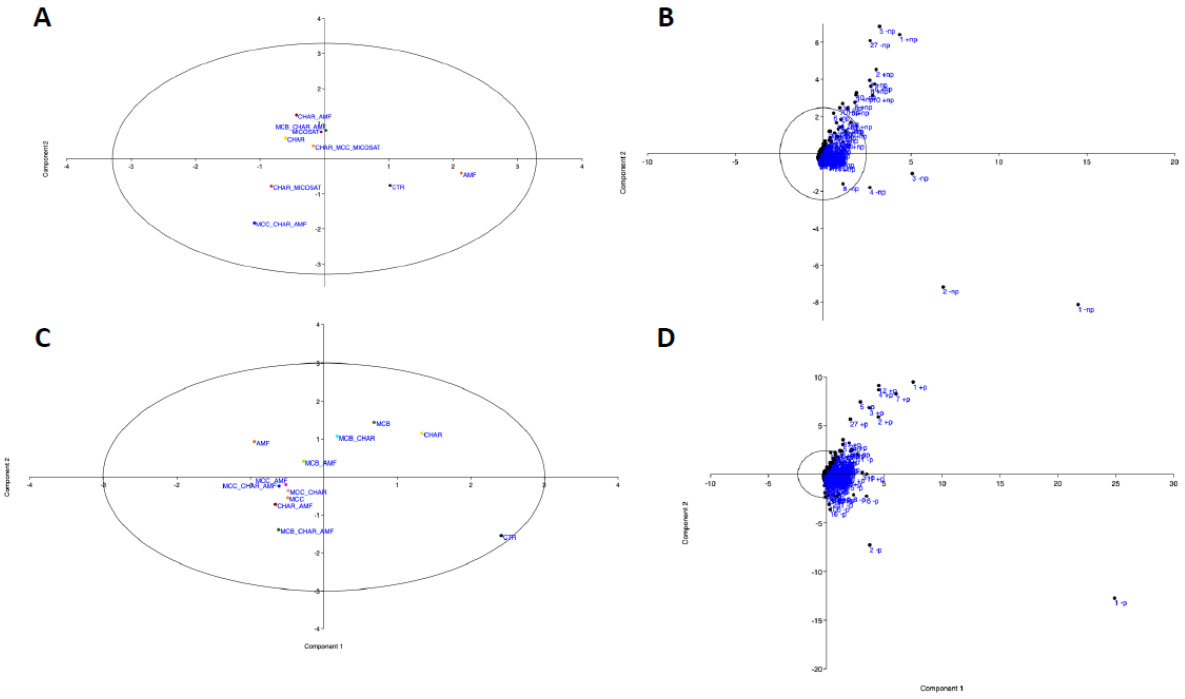


Fig. 4. Principal component analysis (PCA), according to the treatments (A, C) and the metabolites (B, D), of untargeted polar and non polar metabolomes of maize kernels collected after field trails of 2020 (A, B) and 2021 (C, D). (A) Components 1 and 2 explain more than 89% of total variance (74.19% and 15.31%, respectively). (B) Components 1 and 2 explain more than 99% of total variance (98.61% and 0.80%, respectively). (C) Components 1 and 2 explain more than 88% of total variance (83.21% and 5.48%, respectively), while in (D) components 1 and 2 explain almost the total percentage of variance (99.33% and 0.48%, respectively). CONTROL- no addition of biofertilizers or biochar; MC_B and MC_C, synthetic microbial consortia; AMF, mycorrhizae; CHAR, Biochar.

Though it was not possible to obtain a clearcut separation between treatments in 2020, kernels coming from plants supplemented with MC_C (CHAR_MC_C_AMF and CHAR_MC_C_F1) placed a little farther apart from the other samples in the upper left and right sides of the PCA plot (Fig. 4A). PCA

carried out according to metabolites (Fig. 4B), highlighted those responsible for the variance within samples.

The same analysis was performed on the 2021 dataset according to treatments (Fig. 4C) and to metabolites (Fig. 4D): a clear separation of the different samples only for kernels of plants supplemented with MC_B and CHAR located farther along the PC2, indicates metabolomic differences (Fig. 4C).

In 2021, 18 DAMs behaved similarly as compared to 2020 and represent potential novel metabolic biomarkers (Table 1). Many of these were over-accumulated in SynComs treatments as compared to Control: D-pipecolic, Glycerophosphocholine, LysoPE(0:0/18:2(9Z,12Z), PC(18:2(9Z, 12Z)0:0) and eight unidentified species; other species showed an opposite tendency: Glucoheptonic acid, 3E,9Z,12Z-octadecatrienoic acid, and 12-oxo-10E-octadecenoic acid.

4.6. Targeted metabolomics of polar and non-polar compounds

A targeted analysis using a custom metabolomics database was performed on both polar and non-polar species to quantify the level of a set of primary and secondary metabolites with pro-nutritional (vitamins, phenylpropanoids, benzoxazinoids etc) and anti-nutritional (alkaloids, etc) properties (Supplementary Table S10). The above accounted for 70 polar species, including primary and secondary metabolites; and 45 non–polar primary (different lipid classes) and secondary (isoprenoids) metabolites.

In samples of 2020, 51 polar metabolites were quantified. The DAMs were six in CHAR_MC_B_AMF, five in CHAR_MC_C_AMF, and five in CHAR_MC_C_F1, most in the group of phenolic acids, phenylpropanoids and benzoxazinoids (Fig. 5).

In 2021, 31 polar metabolites were quantified; the DAMs were 12 (Supplementary Table S10). Specifically, the classes of amino acids, acids (except for syringic acid) and benzoxazinoids displayed positive alterations, while phenylpropanoids decreased in two samples as in the case of apigenin.

In all the treatments, a significant decrease was observed for azelaic acid, a molecule with anti-inflammatory, antioxidant, and bactericidal effects (Zhou et al., 2018).

Other DAMs increased in 2020 and 2021: rutin and HMBOA-4-Glc over-accumulated in SynComs-treated kernels compared to Control.

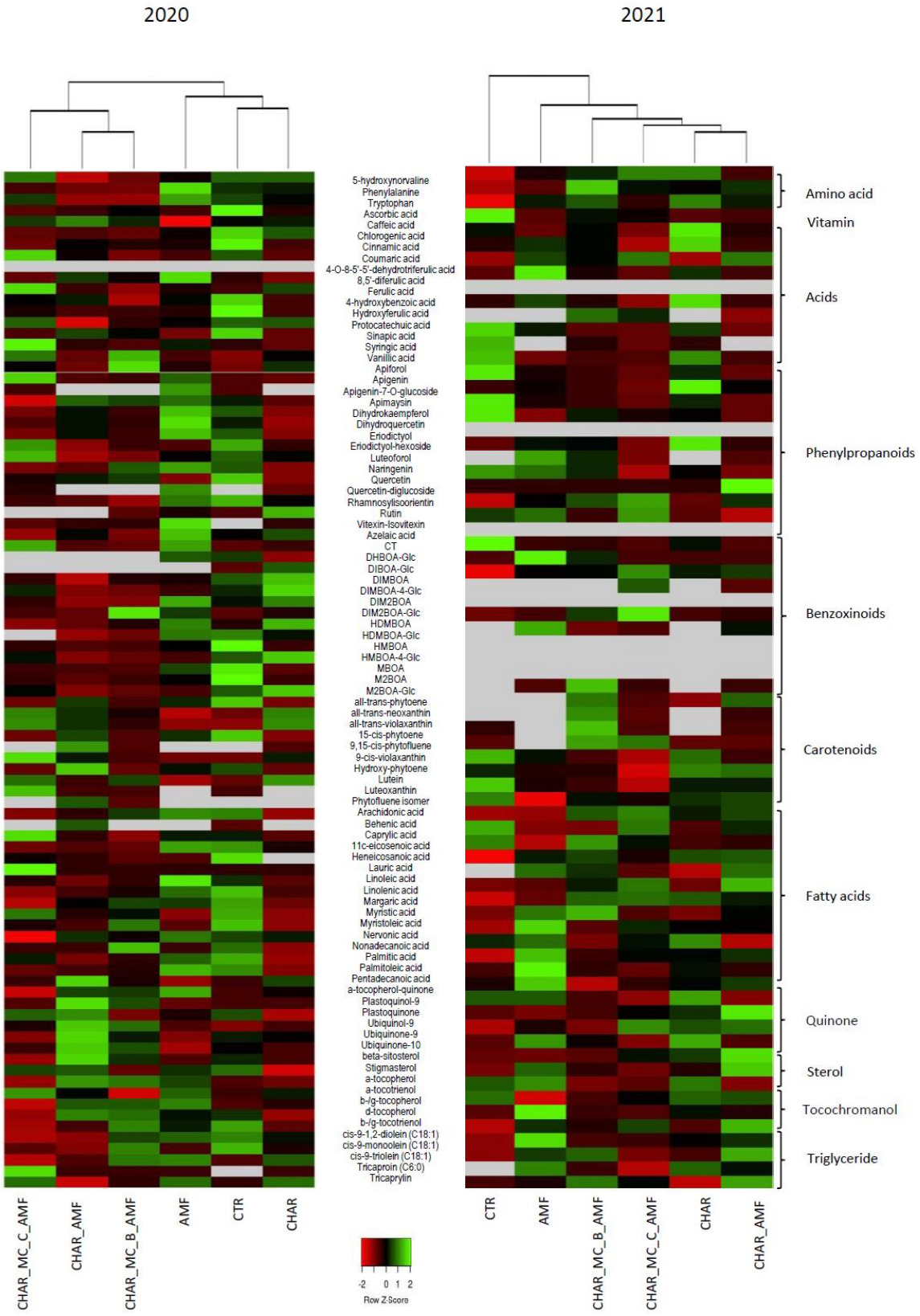


Fig. 5. Heat map representation of HPLC-HRMS targeted metabolomics. Analysis was done of polar and non-polar ionic species detected in kernels of maize cultivated in 2020 and in 2021 from plants grown in presence or absence of microbial consortia B and C, and in combination with AMF, and CHAR. The rows display ionic species and the columns represent the different treatments. colour code scale from green to red indicates the variation in the range from -2 to +2 of down- and up-accumulated metabolites, compared to the control, whereas grey refers to undetected species. CONTROL- no addition of biofertilizers or biochar; MC_B and MC_C, synthetic microbial consortia; AMF, mycorrhizae; CHAR, Biochar.

The targeted analysis of the non-polar metabolome showed specific alterations in treatments with MC_B and MC_C (Supplementary Table S10). In 2020, 45 metabolites were quantified; 10 were differentially abundant in at least one condition as compared to Control: many in the class of triglycerides and fatty acids. The CHAR_MC_B_AMF treatment showed correlations with the highest number of over-abundant compounds.

In 2021, 31 metabolites were quantified, and 12 DAMs differentiated at least one of the treatments with respect to the Control: fatty acids and glycerolipids increased, while isoprenoids decreased. Notably, α−tocotrienol and phytofluene were slightly over-accumulated in treatments with AMF and CHAR_MC_C_AMF, respectively. Linolenic acid, undetected in the Control, was over-accumulated in all the treatments, except for CHAR and MC_C. A comparison of the DAMs from the 2020 and 2021 seasons was performed showing one metabolite, tricaproin (C6:0), characteristic of SynComs-treated vs Control maize kernels.

The set of DAMs that showed over- or down-accumulation in samples from maize grown in the presence of SynComs included both unknown species and annotated molecules from primary and secondary metabolism (Table 1, Fig. 6). For example, D-Pipecolic acid and a group belonging to different lipid classes showed higher content in SynComs-treated vs Control. Interestingly, fatty acids (e.g. 3E,9Z,12Z-octadecatrienoic acid) resulted in down-accumulation in SynComs-samples, whereas phospholipids and diacylglycerols (DGs) showed an opposite tendency. The set DAMs of the two seasons might represent potential metabolic biomarkers of SynComs-plant interactions as well as of crop quality.

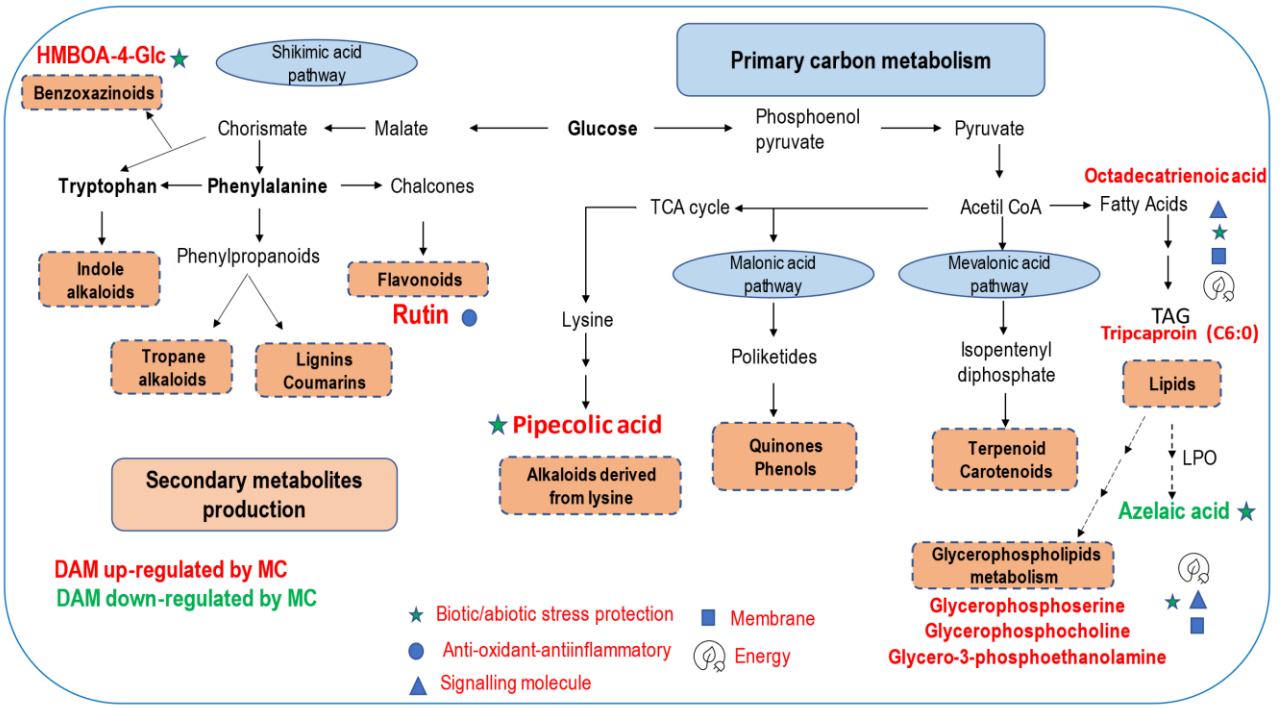


Fig. 6. Overview of the main pathways in primary and secondary metabolism involving the selected set of metabolites. The differentially abundant metabolites (DAM) are reported with the name in red if up-regulated and in green if down-regulated in samples treated with SynComs as compared to the Control.

5. Discussion

5.1. Impacts of biochar and microbial treatments on plant performance

SynComs, alone or in association with biochar, affected plant growth and yield. When comparing sequential seasons, the effects of environmental conditions were more relevant than treatments. In season 2020, SynComs with biochar and mycorrhizae had a significant impact on CHL during vegetative growth; while at maturity, plants grown with biochar had a higher CHL as compared with Control. Also improved was the overall performance of plants grown with CHAR_MC_C_AMF with the highest value of cob FW, plant yield and TKW.

The SIMBA SynComs were also tested on plants cultivated under different fertilization regimes in a different environment (Hett et al., 2023). In 2020 an increase in vegetative growth and nutrient uptake was evident in all inoculated plants, while in 2021 an increased yield was observed only for MC_C_AMF, but not for MC_C. The uneven performance of SynComs might be related to

antagonism between inoculants and native soil microbiota, as well as to abiotic variables: temperature, humidity and soil physic-chemical structure and composition. No beneficial effect under temperate climatic conditions has been reported in wheat inoculated with different *Bacillus* strains (Nguyen, 2018). In field conditions, both biotic and abiotic factors can limit the efficiency of microbial inoculants (Hett et al., 2023; Symanczik et al, 2023), and this explains why field trials can give significant differences from pot experiments.

The combination of SynComs and biochar accounted for most of the variances observed, with the main effects recorded on CHL and on shoots biomass both in previous maize and wheat in pot experiments (Graziano et al., 2022). Field trials in wheat, with some similar combinations of treatments (CHAR and CHAR_F1), provided increased level of CHL (Caldara et al, 2024). Here the same was revealed for maize in 2020.

5.2. Rhizosphere microbial ecology

Testing the union of microbial species with complementary traits was exploited in the last decade due to the advantages over single species (Liu et al., 2023). The plant's rhizosphere selects microorganisms from bulk soil, reducing microbial diversity, but increasing the presence of some classes such as Bacteroidetes, Proteobacteria, and other copiotrophs (Ling et al., 2022). The rhizosphere soil microbiota of maize under the different treatments was relatively stable over the two seasons, with a similar biodiversity level. Furthermore, SynComs application did not significantly change the taxonomic composition of the native rhizosphere microbiota. Differently from previous reports (Cornell et al., 2021) it seems that the complexity and diversity of the rhizosphere ecosystem can mask small changes in species richness (Johnson et al., 2019). Inoculants may increase biodiversity when outcompeting dominant taxa; this may then facilitate competitive release (Cornell et al., 2021; Mawarda et al., 2020). An inoculant may decrease microbial diversity (including increased dominance by some native groups) when its presence increases resource availability or when it becomes dominant and suppresses native taxa (Johnson et al., 2019). A meta-analysis focused

on the effect of bioinoculants showed significant results over the resident soil microbiota diversity in only 62% of analyzed cases (Cornell et al., 2021) and results from field studies frequently disagree with greenhouse studies on the impact of bioinoculants on diversity (Zhao et al., 2022). In the two growing seasons, minor changes were observed (e.g., Actinobacteria were more abundant in 2020 compared to 2021, which could be related to the variation in soil pH). The main phyla were Proteobacteria, Actinobacteria, Chloroflexi, Firmicutes, Verrucomicrobiota and Acidobacteria, key groups of the rhizosphere bacteriome (Berg et al., 2021; Ling et al., 2022; Mishra et al., 2022; van Bergeijk et al., 2020). Studies revealed a shift in soil microbiome after SynComs inoculation although only temporary, as a few weeks after the application the shift is not evident anymore (Berg et al., 2021). In our conditions, the inoculum grew and persisted at least for one month (data not shown). The comparison with pot experiments (Graziano et al., 2022) showed a similar rhizosphere microbial population with no significant modification due to SynComs application, also in combination with biochar.

By comparing wheat and maize field trials under the same conditions (but under different meteorological conditions) it appears that CHAR, F1, and their combination did not significantly alter the bacterial populations (Caldara et al., 2024).

5.3. Kernel metabolites, chemistry, and function as defence, food, feed

Plants counteract environmental stimuli through the production of useful compounds (Medeiros et al., 2021; Nephali et al., 2021). Studies have been conducted on the metabolic reprogramming in the rhizosphere of wheat and sorghum (Carlson et al., 2019; Mashabela et al., 2022), but few reports are available on SynComs-induced metabolic perturbation in maize kernels.

The DAMs detected in the two years showed the presence of glycerophospholipids, which are major constituents of cell membranes (Reszczyńska & Hanaka, 2020), act as signalling molecules (Welti et al., 2007), and are involved in plant defence against pathogens (Cavaco et al., 2021). Fatty acids,

among these Glucoheptonic acid, 3E,9Z,12Z-octadecatrienoic acid, and 12-oxo-10E-octadecenoic acid, were down accumulated during both seasons.

A pathway for the biosynthesis of pipecolic acid was found in several monocot and dicot species (Aliferis et al., 2014; Návarová et al., 2013), including economically and nutritionally important cereals (Garcia-Seco et al., 2017; Moller, 1976). Studies showed that pipecolic acid strongly accumulated in response to phytopathogens in leaves, but also as a compatible solute to protect cells from stressful conditions (Arruda & Barreto, 2020; Kiyota et al., 2015). In the kernel context, its increased amount in conditions of SynComs treatment suggests an attempt to increase resistance to environmental stress.

Kernels from plants treated with SynComs displayed an increased amount of HMBOA-4-Glc, as compared to the control. This species belongs to benzoxazinoids (BXs), indole-derived plant defence compounds (e.g., antifeedant, insecticidal, antimicrobial and allelopathic (Niculaes et al., 2018)) found in wild and cultivated *Poaceae*; and maize is considered a model for studying their role (Kokubo et al., 2017). They are abundant in young tissues in roots and shoots, where they are glucosylated and stored in vacuoles or released as root exudates (Schütz et al., 2019). Lower amounts were detected in wheat and rye seeds, while in maize no clear evidence was found so far (Zhou et al., 2018). In humans, anti-allergenic and anti-inflammatory effects were observed, but it was also reported that BXs suppress appetite and reduce cellular uptake of glucose; furthermore, BXs seem to influence the gut microbial community (Adhikari et al., 2015).

Kernels of SynComs-treated plants displayed over- and down-accumulation of rutin and azelaic acid, respectively. Rutin is a dietary flavonoid found in many vegetables and fruits with therapeutic properties attributed to its antioxidant and anti-inflammatory activities. Studies have demonstrated a broad range of beneficial health properties, like the prevention of neurodegenerative disorders, cardiovascular diseases, and skin cancer, among others. However, health benefits depend on its quantity and bioavailability (Frutos et al., 2019). Azelaic acid, less abundant in kernels of maize

treated with SynComs, is a general marker for lipid peroxidation rather than a general immune signal, and it is accumulated in response to biotic stress (Wu et al., 2013).

Fatty acids, organic acids, and phenylpropanoids were the main classes of DAMs. They can be considered as metabolomic markers for SynComs-induced plant priming and for environmental adaptations to biotic and abiotic stress.

5.4. Carbon Footprint (CF) in transition to biofertilizers

Agriculture accounts for about 30% of GHG (Green House Gases) world emissions, and the four major commodities, maize, wheat, rice, and barley account for about 70% of it. Maize CF was calculated across different environments, and it was estimated in the range of 300-1000 kg CO_2 eq/ton grain (Fantin et al., 2017; Li et al., 2021; Moungsree et al., 2022; Supasri et al., 2020).

When comparing innovative solutions versus conventional cultivation, in the transition towards more sustainable agriculture, the major variables are chemical fertilizers, herbicides, and fungicides. Fertilizers are mainly added as P_2O_5 from sulpho-phosphate at 60 kg/ha, N, from ammonium nitrate and urea at 170 kg/ha, and K at 20 kg/ha (Moungsree et al., 2022). Herbicides and fungicides are given in multiple applications in pre-emergence or post-emergence at ≈ 2-3 L/ha (Fantin et al., 2017). In our experimental plots, N fertilization was reduced by half in 2021, and herbicides/fungicides were omitted.

Chemical fertilizers account for ≈ 80% of the GHG emissions per ton of grains (Fantin et al., 2017; Holka et al., 2017; Moungsree et al., 2022), whereas the contribution to GWP (Global Warming Potential) of all the other phases (sowing, tillage, harvesting, transport) is around 15% (for pesticides ~5%).

Chemistry employed for maize cultivation impacts both the environment (acidification, eutrophication of fresh water, marine eutrophication) and human health (toxicity, carcinogenicity, photochemical ozone formation, resources depletion) (Fantin et al., 2017). A study on corn cultivation in USA (Li et al., 2021) showed a shift of the environmental burdens where ecotoxicity,

carcinogenesis and eutrophication achieved significant values in terms of ton CO_2 eq/ton corn depending on the agricultural practice.

Here, three possible scenarios were designed, to determine CF: 1) conventional cultivation, with fertilizers and pesticides; 2) use of biofertilizers with 50% of N but without pesticides; 3) as in 2) but with biochar (Figure 7A). Scenario 3, with both biofertilizers and biochar, had the lowest CF. Applying similar considerations to the economical input/outputs, for a standard production of about 5-15 ton/ha at a market price of 300 €/ton (please note this figure includes the European subsidies) the economic differences were in favour of the conventional cultivation (Moungsree et al., 2022; Supasri et al., 2020) (Figure 7B, Supplementary Table S11). Therefore, for a standard production between 5-15 ton/ha, there could be a total GWP impact of about 350 kg CO_2 eq/ton grain. Of these emissions, about 60-70% were due to fertilizers and about 5-10% to pesticides and insecticides.

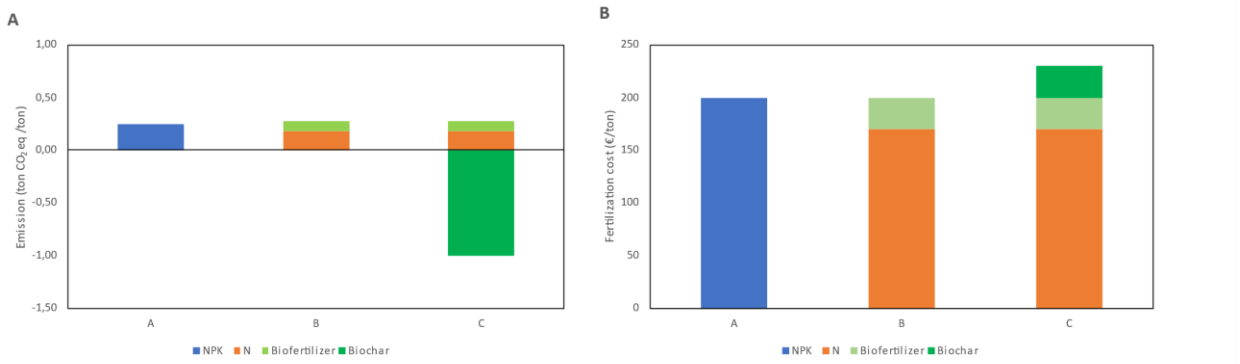


Fig. 7. Impact of the use of biofertilizers and biochar on the fertilization cost (A) and emissions (in terms of ton CO_2 eq/ton grains) as compared to conventional fertilizer. Three scenarios are compared: scenario (A): fertilization with normal NPK fertilization; scenario (B): fertilization with 50 kg/ha N and 15 kg/ha of biofertilizers; scenario (C): fertilization with 50 kg/ha N and 15 kg/ha biofertilizers, with delivery by biochar at 2 ton/ha. This comparison is possible because: the yields between the plots in the three conditions were similar and no disease was observed in plants and grains, and no chemical weeding was necessary. The approach used to discuss in a sustainability perspective the data obtained, is more analytical than empirical, and the data reported are not necessarily related with the input/output of the field experiment but mainly referred from various literature and averaged for simplification.

In terms of Ecological Footprint (EF), the amount of land necessary to regenerate the resources used was estimated at 0.2-0.8 gb ha/ton. Fertilizers account for this in the range between 30 and 40% in conventional agriculture whereas something slightly lower (~1-5%) in the transition to biofertilizers. More difficult was to calculate the benefit of biochar supplements due to a lack of consistency in the literature. Maize has a CO_2 fixation efficiency relatively high (about 80%); higher than soybean but lower than rice and sugarcane (Moungsree et al., 2022). In a USA context, this gives a net profit of 800 USD/ha (about 171 USD/ton of grain) with a net profit (USD/ton CO_2 eq emitted) of 0.44 which strikes when compared with the actual cost of CO_2 emission/ton in the voluntary market ranging from 1 to 100 USD/ton in the absence of governmental support. The situation is strengthened by considering emissions both in terms of GHG and the risk to human and environmental health. The EF of a scenario with a reduced supply of chemical fertilizers seems in favour of these considerations. However, the criticisms on the calculation and use of EF are many and stringent. One for all is that it fails to satisfy simple economic principles and the basic assumptions are theoretically and historically wrong. A piece of this criticism is also evident in our approximate cost calculation.

6. Conclusion

In conclusion, the application of SynComs while not affecting the biodiversity of the native rhizosphere microbial community, thus representing an environmentally friendly tool in view of more sustainable maize cultivation, but did affect kernel quality, with the metabolomic profile of maize kernels enriched in antistress metabolites and health-beneficial molecules. Our results indicate that the contribution of SynComs and biochar could partially replace the intense use of chemical fertilizers. As suggested by LCA studies on maize production, in several countries, major hot spots for all impact categories are fertilizers and pesticides (Boone et al., 2016; Fantin et al., 2017; Holka et al., 2017; Hou et al., 2021; Król-Badziak et al., 2021; Li et al., 2021; Moungsree et al., 2022; Supasri et al., 2020; W. Zhang et al., 2018; Zhong et al., 2020). Transition to more sustainable agriculture is necessary and possible: there are suitable alternatives to chemical fertilizers in terms of

biofertilizers (such as SynComs) and novel amendments (such as biochar) to reduce CF. However, a gross calculation with heterogenous literature data referring to a crop (maize) which has heterogeneous behaviour in different countries, does not show similar economic advantages.

CRediT authorship contribution statement

NM and ABevivino led the project, conceptualized the study and supervised the work. Methodology: MG, MC, SG, LC, SF, GD, AB; AMK. Investigation: MG, SG, MC, LC, SF, AMK. Formal analysis: NM, ABevivino, MG, LC, GD, AF. Software, and Data curation: LC, SF, AMK, SG. Writing – original draft: NM, MG, ABevivino with input from GD, LC, AM. Writing review & editing: NM, MG, ABevivino with input from GD, LC, AM, AB. Visualization: MG, SF, LC. Supervision: NM, ABevivino, GD. Project administration: NM, EM, ABevivino, AP. Funding acquisition, NM, ABevivino, AP. All authors have read and agreed to the presented version of the manuscript.

Acknowledgements

This research was funded by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement N. 818431 (SIMBA, Sustainable Innovation of Microbiome Applications in Food System), grant agreement N. 101112855 (DeliSoil, Delivering safe, sustainable, tailored & societally accepted soil improvers from circular food production processes for boosting soil health), and the Joint Call of the Cofund ERA-Nets SusCrop (grant agreement N. 771134), FACCE ERA-GAS (grant agreement N. 696356), ICT-AGRI-FOOD (grant agreement N. 862665) and SusAn (grant agreement N. 696231). These funding sources had no involvement in study design, collection, analysis and interpretation of data, writing, decision to submit the paper. This output reflects only the authors' views, and the Research Executive Agency (REA) of the EU cannot be held responsible for any use that may be made of the information contained therein. This research was developed in the framework of the European Joint Program for SOIL "Towards climate-smart sustainable management of agricultural soils" (EJP SOIL) funded by the European Union Horizon

2020 research and innovation programme (grant agreement N. 862695) and the Italian project "Creazione di un HUB italiano a supporto della partecipazione dell'Italia alla Global Soil Partnership ed alla rete di eccellenza europea sulla ricerca sul suolo—SOIL-HUB" (D.M. 37072/7303/18—28/12/2018, D.M. 35851—5/112019). The authors gratefully acknowledge: Alfredo Ambrico, Marco Trupo and Alessandra Rosaria Magarelli of ENEA Trisaia Research Center (Rotondella, Italy) and Enrico Ercole from (CCS Aosta Srl, Aosta, Italy) for providing the high-quality microbial inoculants in both years; the personnel at Azienda Stuard (Parma) for technical assistance during the two-year field trials; Giuseppe Aprea, Filippo Sevi, Manuela Costanzo and Silvia Tabacchioni (ENEA Casaccia Research Center) for technical assistance and helpful discussion. This work has benefited from the equipment and framework of the COMP-HUB and COMP-R Initiatives, funded by the 'Departments of Excellence' program of the Italian Ministry for University and Research (MIUR, 2018-2022 and MUR, 2023-2027). The authors thank the European Commission—NextGenerationEU, Project "Strengthening the MIRRI Italian Research Infrastructure for Sustainable Bioscience and Bioeconomy", code n. IR0000005 for the support, and the Microbial Resource Research Infrastructure—Italian Joint Research Unit (MIRRI-IT) for providing strains and data related to ENEA microbial collection.

Supplementary files list:

Supplementary Fig. S1 Agrometeorological conditions recorded during filed trials in seasons 2020 and 2021.

Supplementary Fig. S2. Principal Component Analysis of DeSeq2 microbiota data after a centered log-ratio transformation.

Supplementary Fig. S3. Principal Component Analysis of microbiota with respect to agronomic data.

Supplementary Tables S1. Chemical analysis of soil collected in maize field trials performed in 2020.

Supplementary Tables S2. Chemical analysis of soil collected in maize field trials performed in 2021.

Supplementary Tables S3. Agronomic and physiologic data measured in field trial of season 2020.

Supplementary Tables S4. Agronomic and physiologic data measured in field trial of season 2021.

Supplementary Table S5. Alpha diversity indices.

Supplementary Table S6. Permanova on centered log-ratio transformed of microbiota data.

Supplementary Table S7. Permanova tests performed among treatments at different growth stages and in different years of sampling.

Supplementary Table S8. Untargeted metabolomics dataset of grain samples from maize plants grown in 2020.

Supplementary Table S9. Untargeted metabolomics dataset of grain samples from maize plants in 2021.

Supplementary Table S10. List of targeted analysis of both polar and non-polar metabolites

Supplementary Table S11. Comparison of environmental and economic costs considering three scenarios of fertilization management.

Supplementary Dataset S1. 16S rRNA gene amplicon sequencing. Data analysis and Amplicon Sequence Variants (ASVs) taxonomy.

References

Adhikari, K. B., Tanwir, F., Gregersen, P. L., Steffensen, S. K., Jensen, B. M., Poulsen, L. K., Nielsen, C. H., Høyer, S., Borre, M., & Fomsgaard, I. S. (2015). Benzoxazinoids: Cereal phytochemicals with putative therapeutic and health-protecting properties. In *Molecular Nutrition and Food Research* (Vol. 59, Issue 7, pp. 1324–1338). Wiley-VCH Verlag. https://doi.org/10.1002/mnfr.201400717

Aliferis, K. A., Faubert, D., & Jabaji, S. (2014). A Metabolic Profiling Strategy for the Dissection of Plant Defense against Fungal Pathogens. *PLOS ONE*, *9*(11), e111930. https://doi.org/10.1371/JOURNAL.PONE.0111930

Ammar, E. E., Rady, H. A., Khattab, A. M., Amer, M. H., Mohamed, S. A., Elodamy, N. I., AL-Farga, A., & Aioub, A. A. A. (2023). A comprehensive overview of eco-friendly bio-fertilizers extracted from living organisms. In *Environmental Science and Pollution Research* (Vol. 30, Issue 53, pp. 113119–113137). Springer. https://doi.org/10.1007/s11356-023-30260-x

Arruda, P., & Barreto, P. (2020). Lysine Catabolism Through the Saccharopine Pathway: Enzymes and Intermediates Involved in Plant Responses to Abiotic and Biotic Stress. In *Frontiers in Plant Science* (Vol. 11). Frontiers Media S.A. https://doi.org/10.3389/fpls.2020.00587

Banerjee, S., & van der Heijden, M. G. A. (2022). Soil microbiomes and one health. *Nature Reviews Microbiology 2022 21:1*, *21*(1), 6–20. https://doi.org/10.1038/s41579-022-00779-w

Banerjee, S., Zhao, C., Garland, G., Edlinger, A., García-Palacios, P., Romdhane, S., Degrune, F., Pescador, D. S., Herzog, C., Camuy-Velez, L. A., Bascompte, J., Hallin, S., Philippot, L., Maestre, F. T., Rillig, M. C., & van der Heijden, M. G. A. (2024). Biotic homogenization, lower soil fungal diversity and fewer rare taxa in arable soils across Europe. *Nature Communications 2024 15:1*, *15*(1), 1–10. https://doi.org/10.1038/s41467-023-44073-6

Beesley, L., Marmiroli, M., Pagano, L., Pigoni, V., Fellet, G., Fresno, T., Vamerali, T., Bandiera, M., & Marmiroli, N. (2013). Biochar addition to an arsenic contaminated soil increases arsenic concentrations in the pore water but reduces uptake to tomato plants (Solanum lycopersicum L.). *Science of the Total Environment*, *454–455*, 598–603. https://doi.org/10.1016/j.scitotenv.2013.02.047

Benincasa, P., D'Amato, R., Falcinelli, B., Troni, E., Fontanella, M. C., Frusciante, S., Guiducci, M., Beone, G. M., Businelli, D., & Diretto, G. (2020). Grain Endogenous Selenium and Moderate Salt Stress Work as Synergic Elicitors in the Enrichment of Bioactive Compounds in Maize Sprouts. *Agronomy 2020, Vol. 10, Page 735*, *10*(5), 735. https://doi.org/10.3390/AGRONOMY10050735

Berg, G., Kusstatscher, P., Abdelfattah, A., Cernava, T., & Smalla, K. (2021). Microbiome Modulation—Toward a Better Understanding of Plant Microbiome Response to Microbial Inoculants. *Frontiers in Microbiology*, *12*, 650610. https://doi.org/10.3389/FMICB.2021.650610/BIBTEX

Bindraban, P. S., Dimkpa, C., Nagarajan, L., Roy, A., & Rabbinge, R. (2015). Revisiting fertilisers and fertilisation strategies for improved nutrient uptake by plants. *Biol Fertil Soils*, *51*, 897–911. https://doi.org/10.1007/s00374-015-1039-7

Boone, L., Van linden, V., De Meester, S., Vandecasteele, B., Muylle, H., Roldán-Ruiz, I., Nemecek, T., & Dewulf, J. (2016). Environmental life cycle assessment of grain maize production: An analysis of factors causing variability. *Science of The Total Environment*, *553*, 551–564. https://doi.org/10.1016/J.SCITOTENV.2016.02.089

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods 2016 13:7*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Callens, K., Fontaine, F., Sanz, Y., Bogdanski, A., D'Hondt, K., Lange, L., Smidt, H., van Overbeek, L., Kostic, T., Maguin, E., Meisner, A., Sarand, I., & Sessitsch, A. (2022). Microbiome-based solutions to address new and existing threats to food security, nutrition, health and agrifood systems' sustainability. *Frontiers in Sustainable Food Systems*, *6*, 537. https://doi.org/10.3389/FSUFS.2022.1047765/BIBTEX

Cangioli, L., Mancini, M., Napoli, M., Fagorzi, C., Orlandini, S., Vaccaro, F., & Mengoni, A. (2022). Differential Response of Wheat Rhizosphere Bacterial Community to Plant Variety and Fertilization. *International Journal of Molecular Sciences*, *23*(7). https://doi.org/10.3390/ijms23073616

Carlson, R., Tugizimana, F., Steenkamp, P. A., Dubery, I. A., & Labuschagne, N. (2019). Differential metabolic reprogramming in paenibacillus alvei-primed sorghum bicolor seedlings in response to fusarium pseudograminearum infection. *Metabolites*, *9*(7). https://doi.org/10.3390/metabo9070150

Cavaco, A. R., Ana, ·, Matos, R., & Figueiredo, A. (2021). Speaking the language of lipids: the cross-talk between plants and pathogens in defence and disease. *Cellular and Molecular Life Sciences*, *78*, 4399–4415. https://doi.org/10.1007/s00018-021-03791-0

Chandini, R. K., Kumar, R., Kumar, R., & Om, P. (2019). The Impact of Chemical Fertilizers on our Environment and Ecosystem. In *Research Trends in Environmental Sciences* (2nd Edition, pp. 71–86).

Chaudhary, P., Singh, S., Chaudhary, A., Sharma, A., & Kumar, G. (2022). Overview of biofertilizers in crop production and stress management for sustainable agriculture. *Frontiers in Plant Science*, *13*. https://doi.org/10.3389/FPLS.2022.930340

Chouhan, G. K., Jaiswal, D. K., Gaurav, A. K., Mukherjee, A., & Verma, J. P. (2021). PGPM as a potential bioinoculant for enhancing crop productivity under sustainable agriculture. In *Biofertilizers: Volume 1: Advances in Bio-inoculants* (pp. 221–237). Elsevier. https://doi.org/10.1016/B978-0-12-821667-5.00009-9

Compant, S., Samad, A., Faist, H., & Sessitsch, A. (2019). A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. *Journal of Advanced Research*, *19*, 29–37. https://doi.org/10.1016/J.JARE.2019.03.004

Cornell, C., Kokkoris, V., Richards, A., Horst, C., Rosa, D., Bennett, J. A., & Hart, M. M. (2021). Do Bioinoculants Affect Resident Microbial Communities? A Meta-Analysis. In *Frontiers in Agronomy* (Vol. 3). Frontiers Media S.A. https://doi.org/10.3389/fagro.2021.753474

De Palma, M., Scotti, R., D'Agostino, N., Zaccardelli, M., & Tucci, M. (2022). Phyto-Friendly Soil Bacteria and Fungi Provide Beneficial Outcomes in the Host Plant by Differently Modulating Its Responses through (In)Direct Mechanisms. *Plants*, *11*(20), 2672. https://doi.org/10.3390/PLANTS11202672/S1

de Souza, R. S. C., Armanhi, J. S. L., & Arruda, P. (2020). From Microbiome to Traits: Designing Synthetic Microbial Communities for Improved Crop Resiliency. *Frontiers in Plant Science*, *11*. https://doi.org/10.3389/FPLS.2020.01179

Deshoux, M., Sadet-Bourgeteau, S., Gentil, S., & Prévost-Bouré, N. C. (2023). Effects of biochar on soil microbial communities: A meta-analysis. *Science of The Total Environment*, *902*, 166079. https://doi.org/10.1016/j.scitotenv.2023.166079

Destoumieux-Garzón, D., Mavingui, P., Boetsch, G., Boissier, J., Darriet, F., Duboz, P., Fritsch, C., Giraudoux, P., Roux, F. Le, Morand, S., Paillard, C., Pontier, D., Sueur, C., & Voituron, Y. (2018). The one health concept: 10 years old and a long road ahead. *Frontiers in Veterinary Science*, *5*(FEB), 14. https://doi.org/10.3389/FVETS.2018.00014/BIBTEX

Ding, Y., Liu, Y., Liu, S., Li, Z., Tan, X., Huang, X., Zeng, G., Zhou, L., & Zheng, B. (2016). Biochar to improve soil fertility. A review. *Agronomy for Sustainable Development 2016 36:2*, *36*(2), 1–18. https://doi.org/10.1007/S13593-016-0372-Z

Dono, G., Rambla, J. L., Frusciante, S., Granell, A., Diretto, G., & Mazzucato, A. (2020). Color Mutations Alter the Biochemical Composition in the San Marzano Tomato Fruit. *Metabolites 2020, Vol. 10, Page 110*, *10*(3), 110. https://doi.org/10.3390/METABO10030110

Fantin, V., Righi, S., Rondini, I., & Masoni, P. (2017). Environmental assessment of wheat and maize production in an Italian farmers' cooperative. *Journal of Cleaner Production*, *140*, 631–643. https://doi.org/10.1016/J.JCLEPRO.2016.06.136

FAO, IFAD, UNICEF, WFP, & WHO. (2021). *The State of Food Security and Nutrition in the World 2021. Transforming food systems for food security, improved nutrition and affordable healthy diets for all.* . https://doi.org/10.4060/cb4474en

Fracasso, A., Telò, L., Lanfranco, L., Bonfante, P., & Amaducci, S. (2020). Physiological Beneficial Effect of Rhizophagus intraradices Inoculation on Tomato Plant Yield under Water Deficit Conditions. *Agronomy*, *10*(1). https://doi.org/10.3390/agronomy10010071

Frutos, M. J., Rincón-Frutos, L., & Valero-Cases, E. (2019). Rutin. *Nonvitamin and Nonmineral Nutritional Supplements*, 111–117. https://doi.org/10.1016/B978-0-12-812491-8.00015-1

Garcia-Seco, D., Chiapello, M., Bracale, M., Pesce, C., Bagnaresi, P., Dubois, E., Moulin, L., Vannini, C., & Koebnik, R. (2017). Transcriptome and proteome analysis reveal new insight into proximal and distal responses of wheat to foliar infection by Xanthomonas translucens. *Scientific Reports*, *7*(1), 10157. https://doi.org/10.1038/s41598-017-10568-8

Graziano, S., Caldara, M., Gullì, M., Bevivino, A., Maestri, E., & Marmiroli, N. (2022). A Metagenomic and Gene Expression Analysis in Wheat (T. durum) and Maize (Z. mays) Biofertilized with PGPM and Biochar. *International Journal of Molecular Sciences*, *23*(18), 10376. https://doi.org/10.3390/IJMS231810376/S1

He, M., Xiong, X., Wang, L., Hou, D., Bolan, N. S., Ok, Y. S., Rinklebe, J., & Tsang, D. C. W. (2021). A critical review on performance indicators for evaluating soil biota and soil health of biochar-amended soils. *Journal of Hazardous Materials*, *414*, 125378. https://doi.org/10.1016/j.jhazmat.2021.125378

Herrmann, M. N., Wang, Y., Hartung, J., Hartmann, T., Zhang, W., Nkebiwe, P. M., Chen, X., Müller, T., & Yang, H. (2022). A Global Network Meta-Analysis of the Promotion of Crop Growth, Yield, and Quality by Bioeffectors. *Frontiers in Plant Science*, *13*. https://doi.org/10.3389/fpls.2022.816438

Hett, J., Döring, T. F., Bevivino, A., & Neuhoff, D. (2023). Impact of microbial consortia on organic maize in a temperate climate varies with environment but not with fertilization. *European Journal of Agronomy*, *144*, 126743. https://doi.org/10.1016/J.EJA.2023.126743

Holka, M., Bieńkowski, J. F., Jankowiak, J., & Dąbrowicz, R. (2017). Life cycle assessment of grain maize in intensive, conventional crop production system. *Romanian Agricultural Research*, *34*.

Hou, L., Yang, Y., Zhang, X., & Jiang, C. (2021). Carbon footprint for wheat and maize production modulated by farm size: a study in the North China plain. *International Journal of Climate Change Strategies and Management*, *13*(3), 302–319. https://doi.org/10.1108/IJCCSM-10-2020-0110

Johnson, J. S., Spakowicz, D. J., Hong, B. Y., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E., & Weinstock, G. M. (2019). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-13036-1

Kandlikar, G. S., Gold, Z. J., Cowen, M. C., Meyer, R. S., Freise, A. C., Kraft, N. J. B., Moberg-Parker, J., Sprague, J., Kushner, D. J., & Curd, E. E. (2018). ranacapa: An R package and Shiny web app to explore environmental DNA data with exploratory statistics and interactive visualizations. *F1000Research*, *7*. https://doi.org/10.12688/F1000RESEARCH.16680.1

Kiyota, E., Pena, I. A., & Arruda, P. (2015). The saccharopine pathway in seed development and stress response of maize. *Plant, Cell & Environment*, *38*(11), 2450–2461. https://doi.org/10.1111/PCE.12563

Kokubo, Y., Nishizaka, M., Ube, N., Yabuta, Y., Tebayashi, S. I., Ueno, K., Taketa, S., & Ishihara, A. (2017). Distribution of the tryptophan pathway-derived defensive secondary metabolites gramine and benzoxazinones in Poaceae. *Bioscience, Biotechnology, and Biochemistry*, *81*(3), 431–440. https://doi.org/10.1080/09168451.2016.1256758

Król-Badziak, A., Pishgar-Komleh, S. H., Rozakis, S., & Księżak, J. (2021). Environmental and socio-economic performance of different tillage systems in maize grain production: Application of Life Cycle Assessment and Multi-Criteria Decision Making. *Journal of Cleaner Production*, *278*, 123792. https://doi.org/10.1016/J.JCLEPRO.2020.123792

Li, S., Thompson, M., Moussavi, S., & Dvorak, B. (2021). Life cycle and economic assessment of corn production practices in the western US Corn Belt. *Sustainable Production and Consumption*, *27*, 1762–1774. https://doi.org/10.1016/J.SPC.2021.04.021

Ling, N., Wang, T., & Kuzyakov, Y. (2022). Rhizosphere bacteriome structure and functions. *Nature Communications 2022 13:1*, *13*(1), 1–13. https://doi.org/10.1038/s41467-022-28448-9

Liu, X., Mei, S., & Salles, J. F. (2023). Inoculated microbial consortia perform better than single strains in living soil: A meta-analysis. *Applied Soil Ecology*, *190*, 105011. https://doi.org/10.1016/J.APSOIL.2023.105011

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*, 550. https://doi.org/10.1186/s13059-014-0550-8

Mącik, M., Gryta, A., & Frąc, M. (2020). Biofertilizers in agriculture: An overview on concepts, strategies and effects on soil microorganisms. *Advances in Agronomy*, *162*, 31–87. https://doi.org/10.1016/BS.AGRON.2020.02.001

Maitra, S., Brestic, M., Bhadra, P., Shankar, T., Praharaj, S., Palai, J. B., Shah, M. M. R., Barek, V., Ondrisik, P., Skalický, M., & Hossain, A. (2021). Bioinoculants—Natural Biological Resources for Sustainable Plant Production. *Microorganisms 2022, Vol. 10, Page 51*, *10*(1), 51. https://doi.org/10.3390/MICROORGANISMS10010051

Mapelli, F., Mengoni, A., Riva, V., & Borin, S. (2023). Bacterial culturing is crucial to boost sustainable agriculture. *Trends in Microbiology*, *31*(1), 1–4. https://doi.org/10.1016/J.TIM.2022.10.005

Marmiroli, M., Bonas, U., Imperiale, D., Lencioni, G., Mussi, F., Marmiroli, N., & Maestri, E. (2018). Structural and functional features of chars from different biomasses as potential plant amendments. *Frontiers in Plant Science*, *9*, 1119. https://doi.org/10.3389/FPLS.2018.01119/BIBTEX

Marmiroli, M., Caldara, M., Pantalone, S., Malcevschi, A., Maestri, E., Keller, A. A., & Marmiroli, N. (2022). Building a risk matrix for the safety assessment of wood derived biochars. *Science of The Total Environment*, *839*, 156265. https://doi.org/10.1016/J.SCITOTENV.2022.156265

Mashabela, M. D., Tugizimana, F., Steenkamp, P. A., Piater, L. A., Dubery, I. A., & Mhlongo, M. I. (2022). Untargeted metabolite profiling to elucidate rhizosphere and leaf metabolome changes of wheat cultivars (Triticum aestivum L.) treated with the plant growth-promoting rhizobacteria Paenibacillus alvei (T22) and Bacillus subtilis. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.971836

Mawarda, P. C., Le Roux, X., Dirk van Elsas, J., & Salles, J. F. (2020). Deliberate introduction of invisible invaders: A critical appraisal of the impact of microbial inoculants on soil microbial communities. *Soil Biology and Biochemistry*, *148*, 107874. https://doi.org/10.1016/J.SOILBIO.2020.107874

McMurdie, P. J., & Holmes, S. (2012). Phyloseq: A bioconductor package for handling and analysis of high-throughput phylogenetic sequence data. *Pacific Symposium on Biocomputing*, 235–246. https://doi.org/10.1142/9789814366496_0023

Medeiros, D. B., Brotman, Y., & Fernie, A. R. (2021). The utility of metabolomics as a tool to inform maize biology. In *Plant Communications* (Vol. 2, Issue 4). Cell Press. https://doi.org/10.1016/j.xplc.2021.100187

Mishra, A., Singh, L., & Singh, D. (2022). Unboxing the black box—one step forward to understand the soil microbiome: A systematic review. *Microbial Ecology 2022 85:2*, *85*(2), 669–683. https://doi.org/10.1007/S00248-022-01962-5

Moller, B. L. (1976). Lysine Catabolism in Barley (Hordeum vulgare L.). In *Plant Physiol* (Vol. 57). https://academic.oup.com/plphys/article/57/5/687/6074959

Moungsree, S., Neamhom, T., Polprasert, S., & Patthanaissaranukool, W. (2022). Carbon footprint and life cycle costing of maize production in Thailand with temporal and geographical resolutions. *International Journal of Life Cycle Assessment*, *28*(7), 891–906. https://doi.org/10.1007/S11367-022-02021-4/FIGURES/5

Murtaza, G., Ahmed, Z., Eldin, S. M., Ali, B., Bawazeer, S., Usman, M., Iqbal, R., Neupane, D., Ullah, A., Khan, A., Hassan, M. U., Ali, I., & Tariq, A. (2023). Biochar-Soil-Plant interactions: A cross talk for sustainable agriculture under changing climate. In *Frontiers in Environmental Science* (Vol. 11). Frontiers Media S.A. https://doi.org/10.3389/fenvs.2023.1059449

Návarová, H., Bernsdorff, F., Döring, A. C., & Zeier, J. (2013). Pipecolic Acid, an Endogenous Mediator of Defense Amplification and Priming, Is a Critical Regulator of Inducible Plant Immunity. *The Plant Cell*, *24*(12), 5123–5141. https://doi.org/10.1105/TPC.112.103564

Nephali, L., Moodley, V., Piater, L., Steenkamp, P., Buthelezi, N., Dubery, I., Burgess, K., Huyser, J., & Tugizimana, F. (2021). A Metabolomic Landscape of Maize Plants Treated With a Microbial Biostimulant Under Well-Watered and Drought Conditions. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.676632

Nguyen, M. Luan. (2018). *Biostimulant effects of rhizobacteria on wheat growth and nutrient uptake under contrasted N supplies.* Diss. Université de Liège, Liège, Belgique.

Niculaes, C., Abramov, A., Hannemann, L., & Frey, M. (2018). Plant protection by benzoxazinoids—recent insights into biosynthesis and function. In *Agronomy* (Vol. 8, Issue 8). MDPI AG. https://doi.org/10.3390/agronomy8080143

Nkebiwe, P. M., Weinmann, M., & Müller, T. (2016). Improving fertilizer-depot exploitation and maize growth by inoculation with plant growth-promoting bacteria: From lab to field. *Chemical and Biological Technologies in Agriculture*, *3*(1). https://doi.org/10.1186/s40538-016-0065-5

Nwachukwu, B. C., & Babalola, O. O. (2022). Metagenomics: A Tool for Exploring Key Microbiome With the Potentials for Improving Sustainable Agriculture. In *Frontiers in Sustainable Food Systems* (Vol. 6). Frontiers Media S.A. https://doi.org/10.3389/fsufs.2022.886987

OECD & FAO. (2023). *Environmental sustainability in agriculture 2023*. https://doi.org/10.4060/cc9065en

Pang, Z., Chen, J., Wang, T., Gao, C., Li, Z., Guo, L., Xu, J., & Cheng, Y. (2021). Linking Plant Secondary Metabolites and Plant Microbiomes: A Review. *Frontiers in Plant Science*, *12*, 300. https://doi.org/10.3389/FPLS.2021.621276/BIBTEX

Pavlicevic, M., Abdelraheem, W., Zuverza-Mena, N., O'Keefe, T., Mukhtar, S., Ridge, G., Ranciato, J., Haynes, C., Elmer, W., Pignatello, J., Pagano, L., Caldara, M., Marmiroli, M., Maestri, E., Marmiroli, N., & White, J. C. (2022). Engineered Nanoparticles, Natural Nanoclay and Biochar, as Carriers of Plant-Growth Promoting Bacteria. *Nanomaterials*, *12*(24). https://doi.org/10.3390/nano12244474

Pavlicevic, M., Elmer, W., Zuverza-Mena, N., Abdelraheem, W., Patel, R., Dimkpa, C., O'Keefe, T., Haynes, C. L., Pagano, L., Caldara, M., Marmioli, M., Maestri, E., Marmiroli, N., & White, J. C. (2023). Nanoparticles and biochar with adsorbed plant growth-promoting rhizobacteria alleviate Fusarium wilt damage on tomato and watermelon. *Plant Physiology and Biochemistry*, *203*. https://doi.org/10.1016/j.plaphy.2023.108052

Peng, J., Han, X., Li, N., Chen, K., Yang, J., Zhan, X., Luo, P., & Liu, N. (2021). Combined application of biochar with fertilizer promotes nitrogen uptake in maize by increasing nitrogen retention in soil. *Biochar*, *3*(3), 367–379. https://doi.org/10.1007/s42773-021-00090-6

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), D590–D596. https://doi.org/10.1093/NAR/GKS1219

Rawat, J., Sanwal, P., Saxena, J. (2018). Towards the Mechanisms of Nutrient Solubilization and Fixation in Soil System. In: Meena, V. (eds) Role of Rhizospheric Microbes in Soil. Springer, Singapore. https://doi.org/10.1007/978-981-13-0044-8_8

Reed, L., & Glick, B. R. (2023). The Recent Use of Plant-Growth-Promoting Bacteria to Promote the Growth of Agricultural Food Crops. In *Agriculture (Switzerland)* (Vol. 13, Issue 5). MDPI. https://doi.org/10.3390/agriculture13051089

Reszczyńska, E., & Hanaka, A. (2020). Lipids Composition in Plant Membranes. In *Cell Biochemistry and Biophysics* (Vol. 78, Issue 4, pp. 401–414). Springer. https://doi.org/10.1007/s12013-020-00947-w

Riboni, N., Bianchi, F., Mattarozzi, M., Caldara, M., Gullì, M., Graziano, S., Maestri, E., Marmiroli, N., & Careri, M. (2023). Ultra-high Performance Liquid Chromatography-Ion Mobility-High-Resolution Mass Spectrometry to Evaluate the Metabolomic Response of Durum Wheat to Sustainable Treatments. *Journal of Agricultural and Food Chemistry*, *71*(41), 15407–15416. https://doi.org/10.1021/acs.jafc.3c04532

Riva, V., Mapelli, F., Bagnasco, A., Mengoni, A., & Borin, S. (2022). A Meta-Analysis Approach to Defining the Culturable Core of Plant Endophytic Bacterial Communities. *Applied and Environmental Microbiology*, *88*(6). https://doi.org/10.1128/AEM.02537-21/ASSET/AC6EFC95-29DB-499E-B9A8-25BF0D7776FE/ASSETS/IMAGES/MEDIUM/AEM.02537-21-F008.GIF

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S., Lambin, E., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., De Wit, C. A., Hughes, T., Van Der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., … Walker, B. (2009). *Planetary Boundaries: Exploring the Safe Operating Space for Humanity* (Vol. 14, Issue 2). https://about.jstor.org/terms

Rong, L. bing, Gong, K. yuan, Duan, F. ying, Li, S. kun, Zhao, M., He, J., Zhou, W. bin, & Yu, Q. (2021). Yield gap and resource utilization efficiency of three major food crops in the world – A review. *Journal of Integrative Agriculture*, *20*(2), 349–362. https://doi.org/10.1016/S2095-3119(20)63555-9

Saia, S., Fragasso, M., De Vita, P., & Beleggia, R. (2019). Metabolomics Provides Valuable Insight for the Study of Durum Wheat: A Review. *Journal of Agricultural and Food Chemistry*, *67*(11), 3069–3085. https://doi.org/10.1021/acs.jafc.8b07097

Sánchez-Cañizares, C., Jorrín, B., Poole, P. S., & Tkacz, A. (2017). Understanding the holobiont: the interdependence of plants and their microbiome. *Current Opinion in Microbiology*, *38*, 188–196. https://doi.org/10.1016/J.MIB.2017.07.001

Schmidt, J. E., & Gaudin, A. C. M. (2018). What is the agronomic potential of biofertilizers for maize? A meta-analysis. *FEMS Microbiology Ecology*, *94*(7). https://doi.org/10.1093/femsec/fiy094

Schütz, V., Bigler, L., Girel, S., Laschke, L., Sicker, D., & Schulz, M. (2019). Conversions of benzoxazinoids and downstream metabolites by soil microorganisms. *Frontiers in Ecology and Evolution*, *7*(JUN). https://doi.org/10.3389/fevo.2019.00238

Shayanthan, A., Ordoñez, P. A. C., & Oresnik, I. J. (2022). The Role of Synthetic Microbial Communities (SynCom) in Sustainable Agriculture. *Frontiers in Agronomy*, *4*, 58. https://doi.org/10.3389/FAGRO.2022.896307/BIBTEX

Shetty, S. A., & Lahti, L. (2019). Microbiome data science. *Journal of Biosciences*, *44*(5), 1–6. https://doi.org/10.1007/S12038-019-9930-2/TABLES/1

Siyuan, S., Tong, L., & Liu, R. H. (2018). Corn phytochemicals and their health benefits. In *Food Science and Human Wellness* (Vol. 7, Issue 3, pp. 185–195). Elsevier B.V. https://doi.org/10.1016/j.fshw.2018.09.003

Sohi, S. P., Krull, E., Lopez-Capel, E., & Bol, R. (2010). A Review of Biochar and Its Use and Function in Soil. *Advances in Agronomy*, *105*(1), 47–82. https://doi.org/10.1016/S0065-2113(10)05002-9

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., Biggs, R., Carpenter, S. R., De Vries, W., De Wit, C. A., Folke, C., Gerten, D., Heinke, J., Mace, G. M., Persson, L. M., Ramanathan, V., Reyers, B., & Sörlin, S. (2015). Planetary boundaries: Guiding human development on a changing planet. *Science*, *347*(6223). https://doi.org/10.1126/SCIENCE.1259855/SUPPL_FILE/STEFFEN-SM.PDF

Supasri, T., Itsubo, N., Gheewala, S. H., & Sampattagul, S. (2020). Life cycle assessment of maize cultivation and biomass utilization in northern Thailand. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-60532-2

Symanczik, S., Lipp, C., Mäder, P., Thonar, C., & Kundel, D. (2023). Limited effectiveness of selected bioeffectors combined with recycling phosphorus fertilizers for maize cultivation under Swiss farming conditions. *Frontiers in Plant Science*, *14*. https://doi.org/10.3389/fpls.2023.1239393

Tabacchioni, S., Passato, S., Ambrosino, P., Huang, L., Caldara, M., Cantale, C., Hett, J., Del Fiore, A., Fiore, A., Schlüter, A., Sczyrba, A., Maestri, E., Marmiroli, N., Neuhoff, D., Nesme, J., Sørensen, S. J., Aprea, G., Nobili, C., Presenti, O., … Bevivino, A. (2021). Identification of beneficial microbial consortia and bioactive compounds with potential as plant biostimulants for a sustainable agriculture. *Microorganisms*, *9*(2), 1–23. https://doi.org/10.3390/microorganisms9020426

Takahashi, S., Tomita, J., Nishioka, K., Hisada, T., & Nishijima, M. (2014). Development of a Prokaryotic Universal Primer for Simultaneous Analysis of Bacteria and Archaea Using Next-Generation Sequencing. *PLOS ONE*, *9*(8), e105592. https://doi.org/10.1371/JOURNAL.PONE.0105592

Tan, S., Narayanan, M., Thu Huong, D. T., Ito, N., Unpaprom, Y., Pugazhendhi, A., Lan Chi, N. T., & Liu, J. (2022). A perspective on the interaction between biochar and soil microbes: A way to regain soil eminence. *Environmental Research*, *214*. https://doi.org/10.1016/j.envres.2022.113832

Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., & Polasky, S. (2002). Agricultural sustainability and intensive production practices. *Nature 2002 418:6898*, *418*(6898), 671–677. https://doi.org/10.1038/nature01014

Trivedi, P., Leach, J. E., Tringe, S. G., Sa, T., & Singh, B. K. (2020). Plant–microbiome interactions: from community assembly to plant health. *Nature Reviews Microbiology 2020 18:11*, *18*(11), 607–621. https://doi.org/10.1038/s41579-020-0412-1

Vaccaro, F., Cangioli, L., Mengoni, A., & Fagorzi, C. (2022). Synthetic plant microbiota challenges in nonmodel species. *Trends in Microbiology*, *30*(10), 922–924. https://doi.org/10.1016/j.tim.2022.06.006

van Bergeijk, D. A., Terlouw, B. R., Medema, M. H., & van Wezel, G. P. (2020). Ecology and genomics of Actinobacteria: new concepts for natural product discovery. In *Nature*

Reviews Microbiology (Vol. 18, Issue 10, pp. 546–558). Nature Research. https://doi.org/10.1038/s41579-020-0379-y

Vassura, I., Fabbri, D., Rombolà, A. G., Rizzi, B., Menichetti, A., Cornali, S., Pagano, L., Reggiani, R., Vecchi, M. R., & Marmiroli, N. (2023). Multi-analytical techniques to study changes in carbon and nitrogen forms in a tomato-cultivated soil treated with biochar and biostimulants. *Soil & Environmental Health*, *1*(4), 100050. https://doi.org/10.1016/J.SEH.2023.100050

Vijay, V., Shreedhar, S., Adlak, K., Payyanad, S., Sreedharan, V., Gopi, G., Sophia van der Voort, T., Malarvizhi, P., Yi, S., Gebert, J., & Aravind, P. V. (2021). Review of Large-Scale Biochar Field-Trials for Soil Amendment and the Observed Influences on Crop Yield Variations. In *Frontiers in Energy Research* (Vol. 9). Frontiers Media S.A. https://doi.org/10.3389/fenrg.2021.710766

Welti, R., Shah, J., Li, W., Li, M., Chen, J., Burke, J. J., Fauconnier, M.-L., Chapman, K., Chye, M.-L., & Wang, X. (2007). Plant lipidomics: Discerning biological function by profiling plant complex lipids using mass spectrometry. In *Frontiers in Bioscience* (Vol. 12, Issue 1).

Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*, *3*(2), 180–185. https://doi.org/10.1002/WICS.147

Wright, E. S., Yilmaz, L. S., & Noguera, D. R. (2012). DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. *Applied and Environmental Microbiology*, *78*(3), 717–725. https://doi.org/10.1128/AEM.06516-11

Wu, L., Wang, S., Chen, X., Wang, X., Wu, L., Zu, X., & Chen, Y. (2013). Proteomic and Phytohormone Analysis of the Response of Maize (Zea mays L.) Seedlings to Sugarcane Mosaic Virus. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0070295

Yan, H., Cong, M., Hu, Y., Qiu, C., Yang, Z., Tang, G., Xu, W., Zhu, X., Sun, X., & Jia, H. (2022). Biochar-mediated changes in the microbial communities of rhizosphere soil alter the architecture of maize roots. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/FMICB.2022.1023444/FULL

Zhang, L., Jing, Y., Xiang, Y., Zhang, R., & Lu, H. (2018). Responses of soil microbial community structure changes and activities to biochar addition: A meta-analysis. *Science of The Total Environment*, *643*, 926–935. https://doi.org/10.1016/j.scitotenv.2018.06.231

Zhang, W., Wang, C., Dong, M., Jin, S., & Li, H. (2018). Dynamics of soil fertility and maize growth with lower environment impacts depending on a combination of organic and mineral fertilizer. *Journal of Soil Science and Plant Nutrition*, *18*(2), 556–575. http://data.cma.cn/

Zhao, Y., Lu, G., Jin, X., Wang, Y., Ma, K., Zhang, H., Yan, H., & Zhou, X. (2022). Effects of Microbial Fertilizer on Soil Fertility and Alfalfa Rhizosphere Microbiota in Alpine Grassland. *Agronomy 2022, Vol. 12, Page 1722*, *12*(7), 1722. https://doi.org/10.3390/AGRONOMY12071722

Zhong, F., Jiang, D., Zhao, Q., Guo, A., Ullah, A., Yang, X., Cheng, Q., Zhang, Y., & Ding, X. (2020). *Eco-efficiency of oasis seed maize production in an arid region, Northwest China*. https://doi.org/10.1016/j.jclepro.2020.122220

Zhou, S., Richter, A., & Jander, G. (2018). Beyond defense: Multiple functions of benzoxazinoids in maize metabolism. In *Plant and Cell Physiology* (Vol. 59, Issue 8, pp. 1528–1533). Oxford University Press. https://doi.org/10.1093/pcp/pcy064

Table 1 List of differentially accumulated metabolites identified as the potentially metabolic biomarkers for SynComs-plant interactions and crop quality. Data are coming from both untargeted and targeted metabolomics analyses of maize kernel samples collected in field trials of seasons 2020 and 2021.

m/z	Fraction	Voltage	Metabolite ID	2020	2021
130.0864	Polar	pos	D-Pipecolic acid	over-	over-
267.0130	Polar	pos	Nf	over-	over-
279.2318	Polar	pos	3E,9Z,12Z-octadecatrienoic acid	down-	down-
520.3399	Polar	pos	PC(18:2(9Z,12Z)/0:0) 1-linoleoyl-2-(1-enyl-stearoyl)-sn-glycero-3-phosphocholine	over-	over-
521.3432	Polar	pos	Nf	over-	over-
564.3307	Polar	neg	Nf	over-	over-
565.3340	Polar	neg	Nf	over-	over-
296.0659	Polar	pos	Glycerophosphocholine	over-	over-
476.2783	Polar	neg	LysoPE(0:0/18:2(9Z,12Z)) 2-linoleoyl-sn-glycero-3-phosphoethanolamine (glycerophospholipid)	over-	over-
522.3558	Polar	pos	PC(O-16:1(11Z)/2:0) 1-(11Z-hexadecenyl)-2-acetyl-sn-glycero-3-phosphocholine (Glycerophospholipids)	over-	over-
566.3466	Polar	neg	PS (21:0/0:0) 1-heneicosanoyl-glycero-3-phosphoserine (Glycerophospholipids)	over-	over-
891.2847	Polar	neg	Nf	over-	over-
576.5063	Non polar	pos	Nf	over-	over-
187.0975	Polar	neg	azelaic acid	down-	down-
387.2740	Non polar	pos	tricaproin (C6:0)	over-	over-
358.1132	Polar	pos	HMBOA-4-Glc	over-	over-
611.1606	Polar	pos	rutin	over-	over-

AMERICAN SOCIETY FOR MICROBIOLOGY | mSystems®

RESEARCH ARTICLE
Host-Microbe Biology

Nonadditive Transcriptomic Signatures of Genotype-by-Genotype Interactions during the Initiation of Plant-Rhizobium Symbiosis

Camilla Fagorzi,[a] Giovanni Bacci,[a] Rui Huang,[b] Lisa Cangioli,[a] Alice Checcucci,[a]* Margherita Fini,[a] Elena Perrin,[a] Chiara Natali,[a] George Colin diCenzo,[b] Alessio Mengoni[a]

[a]Department of Biology, University of Florence, Florence, Italy
[b]Department of Biology, Queen's University, Kingston, Ontario, Canada

ABSTRACT Rhizobia are ecologically important, facultative plant-symbiotic microbes. In nature, there is a large variability in the association of rhizobial strains and host plants of the same species. Here, we evaluated whether plant and rhizobial genotypes influence the initial transcriptional response of rhizobium following perception of a host plant. RNA sequencing of the model rhizobium *Sinorhizobium meliloti* exposed to root exudates or luteolin (an inducer of *nod* genes, involved in the early steps of symbiotic interaction) was performed on a combination of three *S. meliloti* strains and three alfalfa varieties as host plants. The response to root exudates involved hundreds of changes in the rhizobium transcriptome. Of the differentially expressed genes, 35% were influenced by the strain genotype, 16% were influenced by the plant genotype, and 29% were influenced by strain-by-host plant genotype interactions. We also examined the response of a hybrid *S. meliloti* strain in which the symbiotic megaplasmid (~20% of the genome) was mobilized between two of the above-mentioned strains. Dozens of genes were upregulated in the hybrid strain, indicative of nonadditive variation in the transcriptome. In conclusion, this study demonstrated that transcriptional responses of rhizobia upon perception of legumes are influenced by the genotypes of both symbiotic partners and their interaction, suggesting a wide spectrum of genetic determinants involved in the phenotypic variation of plant-rhizobium symbiosis.

IMPORTANCE A sustainable way for meeting the need of an increased global food demand should be based on a holobiont perspective, viewing crop plants as intimately associated with their microbiome, which helps improve plant nutrition, tolerance to pests, and adverse climate conditions. However, the genetic repertoire needed for efficient association with plants by the microbial symbionts is still poorly understood. The rhizobia are an exemplary model of facultative plant symbiotic microbes. Here, we evaluated whether genotype-by-genotype interactions could be identified in the initial transcriptional response of rhizobium perception of a host plant. We performed an RNA sequencing study to analyze the transcriptomes of different rhizobial strains elicited by root exudates of three alfalfa varieties as a proxy of an early step of the symbiotic interaction. The results indicated strain- and plant variety-dependent variability in the observed transcriptional changes, providing fundamentally novel insights into the genetic basis of rhizobium-plant interactions. Our results provide genetic insights and perspective to aid in the exploitation of natural rhizobium variation for improvement of legume growth in agricultural ecosystems.

KEYWORDS RNA-seq, *Sinorhizobium meliloti*, hybrid strain, plant-microbe interactions

Microbes play a crucial role in the biology and evolution of their eukaryotic hosts (1). Among other activities, microbes contribute to the host's acquisition of

Citation Fagorzi C, Bacci G, Huang R, Cangioli L, Checcucci A, Fini M, Perrin E, Natali C, diCenzo GC, Mengoni A. 2021. Nonadditive transcriptomic signatures of genotype-by-genotype interactions during the initiation of plant-rhizobium symbiosis. mSystems 6:e00974-20. https://doi.org/10.1128/mSystems.00974-20.

Editor Davide Bulgarelli, University of Dundee

Copyright © 2021 Fagorzi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Giovanni Bacci, giovanni.bacci@unifi.it, or Alessio Mengoni, alessio.mengoni@unifi.it.

* Present address: Alice Checcucci, Department of Agricultural and Food Science, University of Bologna, Bologna, Italy.

Symbiont tasting: transcriptomic signatures of host plant variety perception by rhizobia

Received 2 October 2020
Accepted 21 December 2020
Published 12 January 2021

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

nutrients (2), functioning of the host's immune system (3), and protection of the host from predation (4). The rules governing host-microbe interactions remain a topic of intense investigation. In many cases, the eukaryotic host selectively recruits the desired microbial partner: squid light organs are selectively colonized by *Vibrio* symbionts (5), legumes select for effective symbionts by sanctioning noneffective symbionts (6), and the crop microbiome is cultivar dependent (7, 8). The genetic basis determining the quality of a microbial symbiont (i.e., its ability to improve host plant phenotypes such as growth and tolerance) and its ability to effectively colonize its eukaryotic partner is generally not well understood, but evolution experiments and high-throughput genome sequencing projects of host-associated microbes and complete microbiomes are shedding light on this topic (9–14). In the case of plants, such studies have observed an enrichment of certain gene functions in plant-associated microbes, such as genes related to carbohydrate metabolism, secretion systems, phytohormone production, and phosphorus solubilization (11, 12, 15, 16).

The rhizobia are an ecologically important exemplar of facultative host-associated microbes. These soil-dwelling bacteria are able to colonize plants and enter an endosymbiotic association with plants of the family *Fabaceae* (17). This developmentally complex process begins with an exchange of signals between the free-living organisms (18), which leads to the invasion of the plant by the rhizobia (19), and culminates in the formation of a new organ (a nodule) in which the plant cells are intracellularly colonized by N_2-fixing rhizobia (20, 21). Decades of research have identified an intricate network of coordinated gene functions required to establish a successful mutualistic interaction between rhizobia and legumes (21–23). In contrast to the core symbiotic machinery, most of which has been elucidated, much remains unknown about the accessory genes required to optimize the interaction.

In addition to simple gene presence/absence, genotype-by-genotype (GxG) interactions have prominent impacts on symbiotic outcomes (24). The importance of both the plant and bacterial genotypes, and their interaction, in optimizing symbioses between rhizobia and legumes was recognized in early population genetic studies (25–27). More recently, greenhouse studies have directly demonstrated the influence of GxG interactions on the fitness of both the plant and rhizobium partners (28–31). The newly developed select-and-resequence approach is providing a high-throughput approach to uncover the genetic basis underlying GxG interactions for fitness in rhizobium-legume symbioses as well as a way to screen for strain-specific effects of individual genes (32, 33). To date, GxG interaction studies have largely focused on measurements of fitness as a holistic measure of the entire symbiotic process. Nodule formation is a complex developmental process involving several steps, each of which requires a distinct molecular toolkit (34), and in principle, distinct GxG interactions could be acting at each of these developmental stages. Transcriptomic studies have demonstrated that GxG interactions have significant impacts on the gene expression patterns of both partners in mature N_2-fixing nodules (35, 36). However, we are unaware of studies specifically focusing on the role of GxG interactions in early developmental stages, such as during the initial perception of the partners by each other. Such knowledge is critical not only to fully understand the microevolution of host-associated bacteria but also to develop host variety-specific rhizobium bioinoculants that may ensure good nodulation abilities over unwanted (indigenous) rhizobial strains (37, 38).

Here, we evaluated whether GxG interactions could be identified in the initial transcriptional response of rhizobium perception of a host plant. We worked with *Sinorhizobium meliloti*, which is one of the best-studied models for GxG interactions in rhizobia. *S. meliloti* forms N_2-fixing nodules on plants belonging to the tribe *Trigonelleae* (39), which includes alfalfa, a major forage crop grown worldwide for which many varieties have been developed (40). The *S. meliloti* genome comprises three main replicons, a chromosome, a chromid, and a megaplasmid; the latter one harbors most of the essential symbiotic functions, including the genes responsible for the initial molecular dialog with the

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

host plant (*nod* genes) (41, 42). To address our aim, the gene expression patterns of three strains of *S. meliloti* (each with distinct symbiotic properties) following 4 h of exposure to root exudates derived from three alfalfa cultivated varieties were characterized using RNA sequencing (RNA-seq). The transcriptome following exposure to luteolin (a known inducer of *nod* genes, involved in the early steps of symbiotic interaction [43]) was also analyzed. Additionally, the relevance of the megaplasmid in defining the strain-specific transcriptional responses was analyzed by studying a hybrid *S. meliloti* strain in which the native megaplasmid was replaced with that of another wild-type strain. The results demonstrated that the transcriptional response involved genes on all three replicons and that, even among conserved *S. meliloti* genes, transcriptional patterns were both strain and root exudate specific.

RESULTS

Symbiotic phenotypes differ across rhizobial strain-plant variety combinations. Symbiotic phenotypes (plant growth and nodule number) and root adhesion of *S. meliloti* strains Rm1021, BL225C, and AK83 were measured during interactions with three varieties of alfalfa (Camporegio, Verbena, and Lodi). The results indicated that these phenotypes are influenced by both the plant and bacterial genotypes (Fig. 1; see also Fig. S1 in the supplemental material). Root adhesion phenotypes (Fig. 1a) were divided by the Scott-Knott test into three main groups reflecting high, medium, and low root colonization. Interestingly, each group was heterogeneous with respect to both plant variety and *S. meliloti* strain, consistent with the specificity of plant variety (i.e., genotype *sensu lato*) and strain individuality (i.e., strain genotype) pairs in root colonization efficiency. For instance, *S. meliloti* BL225C strongly colonized the roots of the Camporegio and Verbena varieties, but it displayed much weaker colonization of the Lodi cultivar. On the other hand, *S. meliloti* AK83 colonized the Lodi and Camporegio varieties better than the Verbena cultivar. Nodules per plant as well as measures of symbiotic efficiency (epicotyl length and shoot dry weight) showed differences among the strain-variety combinations (Fig. 1b to d). However, the extents of the variation were lower than those recorded for plant root adhesion. The highest number of nodules was found on the Lodi variety nodulated by *S. meliloti* AK83, which was previously interpreted as a consequence of its reduced N_2 fixation ability with some alfalfa varieties (44–46). Interestingly, the measures of symbiotic efficiency did not correlate with root adhesion phenotypes (both adhesion versus dry weight and adhesion versus epicotyl length gave nonsignificant Pearson correlation values [$P > 0.18$]). However, we cannot *a priori* exclude that measuring adhesion over the whole root might not reflect adhesion to the root hair extension zone, where rhizobia start the symbiotic interaction. For example, the largest plants were the Lodi variety inoculated with *S. meliloti* BL225C despite the root adhesion of this combination being the lowest. Similarly, the smallest plants were the Verbena variety inoculated with *S. meliloti* Rm1021 despite strong root adhesion in this pairing.

Root exudates differ among alfalfa varieties. Liquid chromatography-mass spectrometry (LC-MS) analysis of the alfalfa root exudates detected a total of 2,688 unique features, including 392 annotated features, across the two platforms: 1,514 hydrophilic features were detected by ultraperformance liquid chromatography (UPLC)-MS in positive mode (PP) (288 annotated), and 1,174 hydrophilic features were detected by UPLC-MS in negative mode (PN) (104 annotated) (see worksheet 1 in Data Set S1 in the supplemental material). In order to clarify if the metabolite compositions of the root exudates from the alfalfa varieties differed, principal-component analysis (PCA) was performed on the two biological replicates of the three cultivars (Fig. 2). The three cultivars clearly grouped separately from each other, suggesting the presence of variety-specific differences in their metabolic compositions. Peaks PP_23583, PP_25608, PP_14051, and PP_23300 were assigned by the PubChem database to liquiritigenin, apigenin, genistein, and apigeninidin, respectively; however, differences in the concentrations of these compounds between the root exudates were not statistically significant (data not shown). Most of the observed differences were related to amino acids, in particular *N*-acetyl-L-leucine, tryptophan, cytosine, 3,5-dihydroxyphenylglycine, and

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

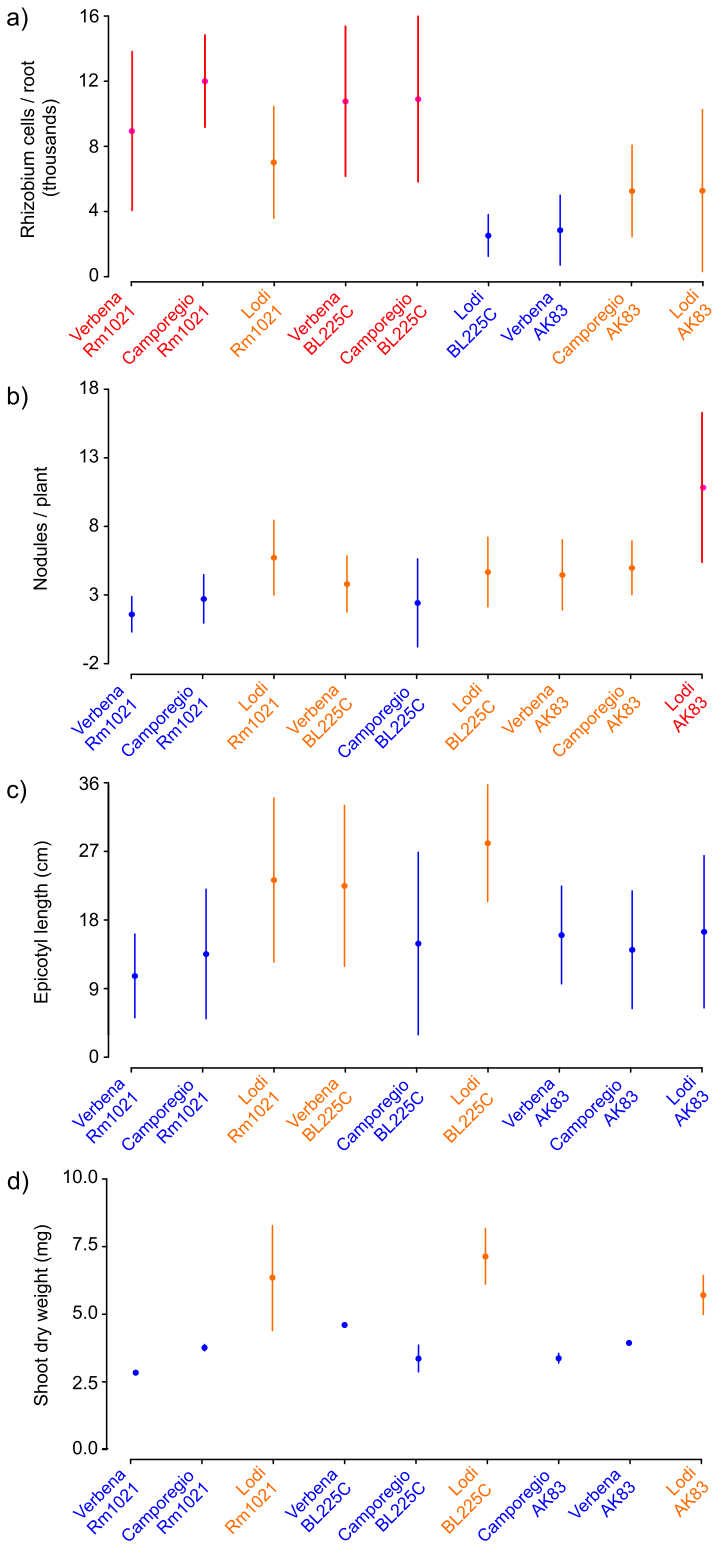


FIG 1 Strain-by-plant variation of symbiosis-associated phenotypes. The number of rhizobium cells retrieved from plant roots (a), number of nodules per plant (b), epicotyl length (c), and plant dry weight (d) are reported. Different colors (pink, orange, and blue) indicate statistically significant groupings ($P < 0.05$) based on a Scott-Knott test. For each condition, the dots indicate the mean values, and the vertical lines indicate the standard deviations.

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

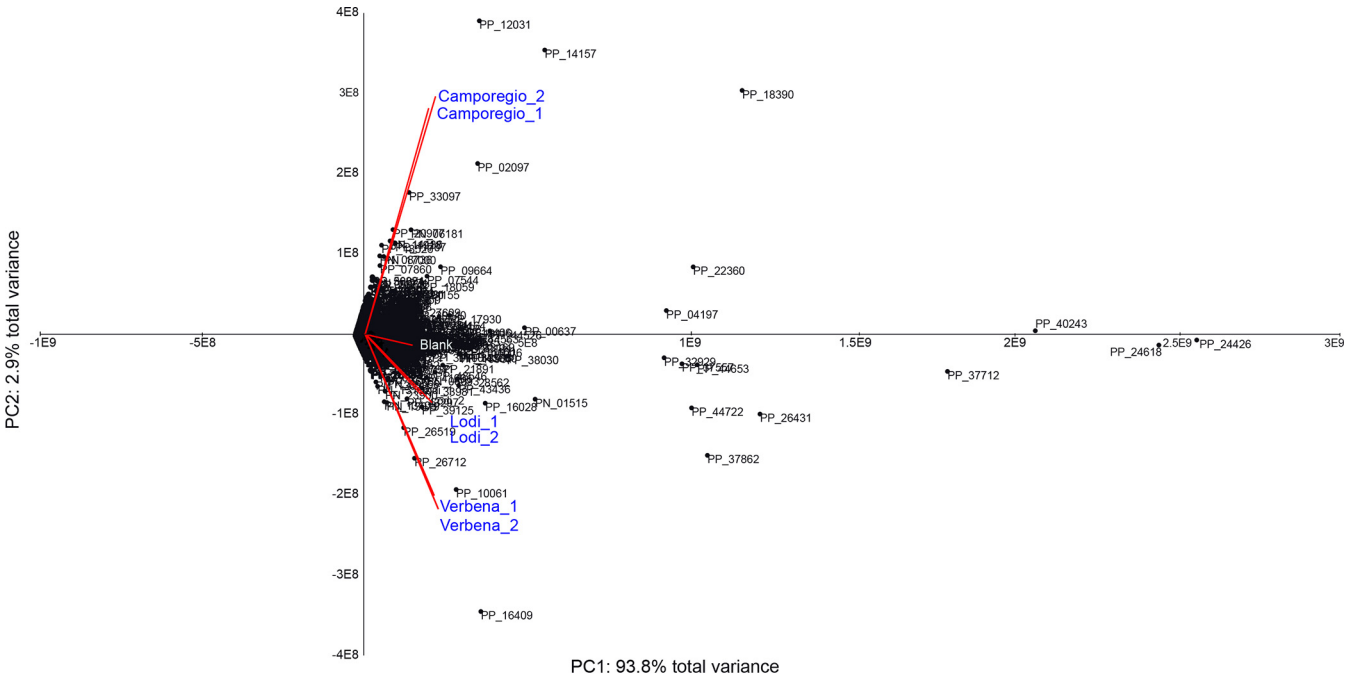


FIG 2 Principal-component analysis biplot from LC-MS analysis of root exudates of the Verbena, Lodi, and Camporegio varieties of alfalfa, including the blank control. Centroids report LC-MS peak IDs (see worksheet 1 in Data Set S1 in the supplemental material), and vectors indicate the loadings of plant varieties.

the dipeptide Val-Ala (Table S1). Multiple flavones and flavonoids, which include known inducers of NodD activation (43) and chemotaxis (47), were potentially identified. These include a peak hypothetically attributed to apigeninidin (PP_23300), which was found in the Verbena and Camporegio root exudates; liquiritigenin (PP_23583), which was found in the Camporegio and Lodi root exudates; as well as apigenin (PP_25608) and genistein (PP_14051), which were found in variable amounts in the root exudates from all three varieties. Elemental analysis (CHNS [carbon, hydrogen, nitrogen, sulfur]) of root exudates was also performed (Table S2), and the results were used to normalize the quantity of root exudates used in the treatment of *S. meliloti* strains based on equalizing the amount of total organic carbon (TOC) added to each culture.

The number of differentially expressed genes changes in strain-condition combinations. The global transcriptional responses of the three *S. meliloti* wild-type strains following a 4-h exposure to luteolin (the model flavone involved in the early steps of symbiotic interaction [43]) or alfalfa root exudates were evaluated using RNA sequencing. In addition, a fourth strain (BM806, referred to as "hybrid" for simplicity) was included (48); the results for this strain are discussed below. A list of differentially expressed genes (DEGs) for each strain and condition (luteolin and the root exudates of the three plant varieties) against the control (blank sample) is reported in Data Set S1, worksheet 2. DEGs were considered to be biologically significant if they had a $\geq$2-fold change in expression and an adjusted *P* value of $<$0.01. The numbers of DEGs are shown in Table 1 (also see Data Set S1, worksheet 3). Reverse transcriptase quantitative PCR (RT-qPCR) on a panel of seven DEGs validated the reliability of the RNA-seq data (Table S3).

In general, luteolin treatment resulted in the lowest number of DEGs, ranging from 36 to 149 per strain. Concerning the root exudates, the number of DEGs was influenced by both the strain and the alfalfa cultivar. Overall, the Camporegio and Verbena root exudates induced more gene expression changes than the Lodi root exudate. Cluster analyses of all genes that were differentially expressed under at least one condition (fold change of $\geq$2; adjusted *P* value of $<$0.01) revealed that for each strain, the

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

TABLE 1 Significant DEGs[a]

Strain	No. of significant DEGs (%)			
	Camporegio	Lodi	Verbena	Luteolin
1021	516 (8.79)	32 (0.55)	506 (8.62)	36 (0.61)
AK83	357 (5.84)	66 (1.08)	192 (3.14)	60 (0.98)
BL225C	1,159 (19.33)	76 (1.27)	693 (11.56)	149 (2.49)
Hybrid	503 (8.38)	98 (1.63)	325 (5.41)	52 (0.87)

[a]The number of significant DEGs with respect to the blank control (2-fold change in expression and an adjusted *P* value of ≤0.01) and the percentage with respect to the total number of genes are reported.

transcriptional responses to the Verbena and Camporegio root exudates were similar and grouped separately from that of the Lodi cultivar (Fig. 3a; Fig. S2 [see also supplemental File S1 at https://doi.org/10.5061/dryad.jdfn2z38q]). Interestingly, ~80% of the genes upregulated by root exudates were found on the chromosomes of the three strains, whereas ~77% of the downregulated genes were found on the pSymA and pSymB replicons (Data Set S1, worksheet 3). This is consistent with a previous signature-tagged mutagenesis study reporting that 80% of genes required for rhizosphere colonization are chromosomally located in *S. meliloti* Rm1021 (49).

Under all conditions, *S. meliloti* BL225C displayed the largest number of DEGs (with up to 20% of genes differentially expressed) (Fig. 4e and f), while *S. meliloti* AK83 had the fewest (Fig. 4c and d). The majority of DEGs (>75%) had orthologs in all three of the tested strains (Data Set S1, worksheet 4), Interestingly, ≥90% of genes upregulated in response to root exudate exposure belonged to the core genome of the three *S. meliloti* strains (Data Set S1, worksheet 2), suggesting that the large majority of genes required for alfalfa rhizosphere colonization are highly conserved. However, expression patterns were not necessarily conserved, and strain-by-strain and condition-dependent variability of the expression pattern on the conserved gene set was observed (Fig. 4; Fig. S3). Indeed, nested likelihood ratio tests (LRTs) indicated that up to 29% of the conserved genes were influenced by strain-condition interactions, consistent with an important role of GxG interactions in the initiation of rhizobium-legume symbioses (Table 2). Moreover, the same analysis emphasized the role of strain genotype in the response to a common condition (35% of associated DEGs).

Stimulons differ in the set of elicited functions. Functional enrichment analyses, based on Kyoto Encyclopedia of Genes and Genomes (KEGG) modules and Clusters of Orthologous Genes (COG) categories, were performed to give a global overview of the functions of the DEGs (Table 3 [see also File S2 at https://doi.org/10.5061/dryad.jdfn2z38q]). Strain- and condition-specific patterns of functional enrichment were observed, consistent with the functional differentiation of the stimulons from each experiment. Nevertheless, a core set of COG categories were commonly over- or underrepresented in all three *S. meliloti* strains during exposure to the Camporegio or Verbena root exudates. These included enrichment among the upregulated genes of COG categories J and O related to protein expression and modification, suggesting that the root exudates stimulated major remodeling of the proteome. In addition, for upregulated genes, COG category G (carbohydrate transport and metabolism) was underrepresented, while for the downregulated genes, COG category C (energy production and conversion) was overrepresented. This observation suggests that the root exudates stimulated a global change in the cellular energy production pathways versus growth in our standard minimal medium with succinate as the sole carbon source. A comparison with growth under soil-mimicking conditions would be interesting with respect to interpreting the root exudate-induced changes in an ecological context.

Among the most highly expressed genes in *S. meliloti* Rm1021 during exposure to the Verbena and Camporegio root exudates were *smc03024* and *smc03028*, encoding components of the flagellar apparatus (*flgF* and *flgC*, respectively); the orthologs of these genes were not induced in BL225C or AK83 (Data Set S1, worksheet 2). The

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

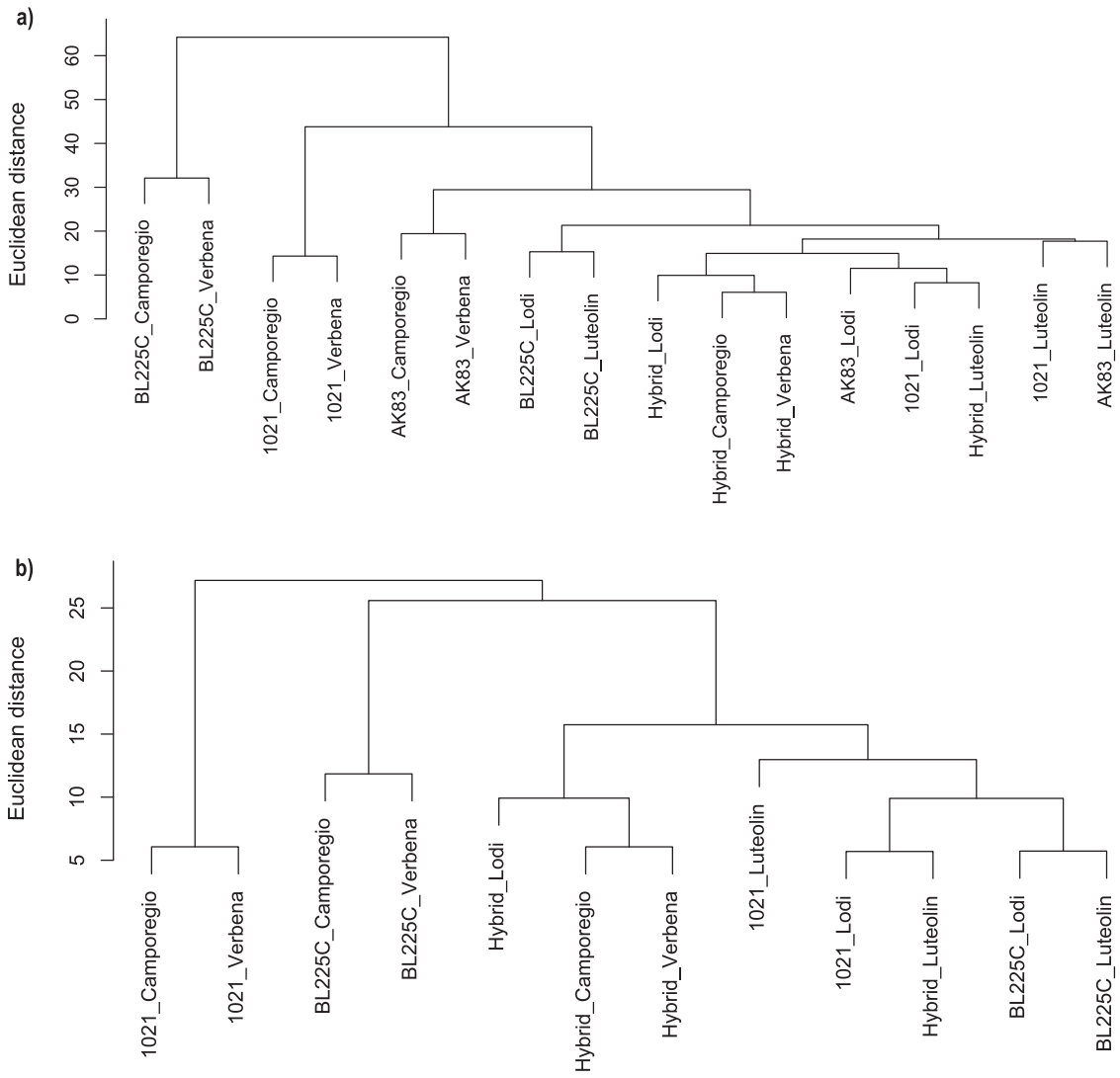


FIG 3 Cluster analyses of the expression profiles of the conserved gene sets of strains grown in the presence of the three root exudates and luteolin. (a) Full data set of DEGs identified in the four strains. (b) DEGs of orthologs of pSymA-like only. Hierarchical clustering was performed in R with the hclust function based on Euclidean distance, with the "complete" agglomeration method. Bars represent Euclidean distance.

induction of motility is in contrast to the observation that luteolin alone decreases the motility of the *S. meliloti* Rm1021 strain (50, 51). Presumably, this reflects the presence of additional stimuli in the root exudates. Indeed, amino acids present in root exudates are known to stimulate chemotactic behavior in *S. meliloti* (52), and signature-tagged mutagenesis showed that motility-related genes are relevant during competition for rhizosphere colonization by *S. meliloti* Rm1021 (49).

Differences in the transcriptomes of two *Bradyrhizobium diazoefficiens* strains exposed to root exudates were suggested to be related to differences in their competitive abilities (53). We therefore examined the expression patterns of several genes likely to play a role in competition for rhizosphere colonization and root adhesion. It was previously suggested that the *sin* quorum sensing system is involved in competition in *S. meliloti* (54); in our data, *sinI* (*smc00168*) was repressed in *S. meliloti* Rm1021 in the presence of the Camporegio and Verbena root exudates, but no changes in the expression of the orthologous genes in strain AK83 or BL225C were observed. No evidence was found in any of the strains for changes in the expression of galactoglucan or

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

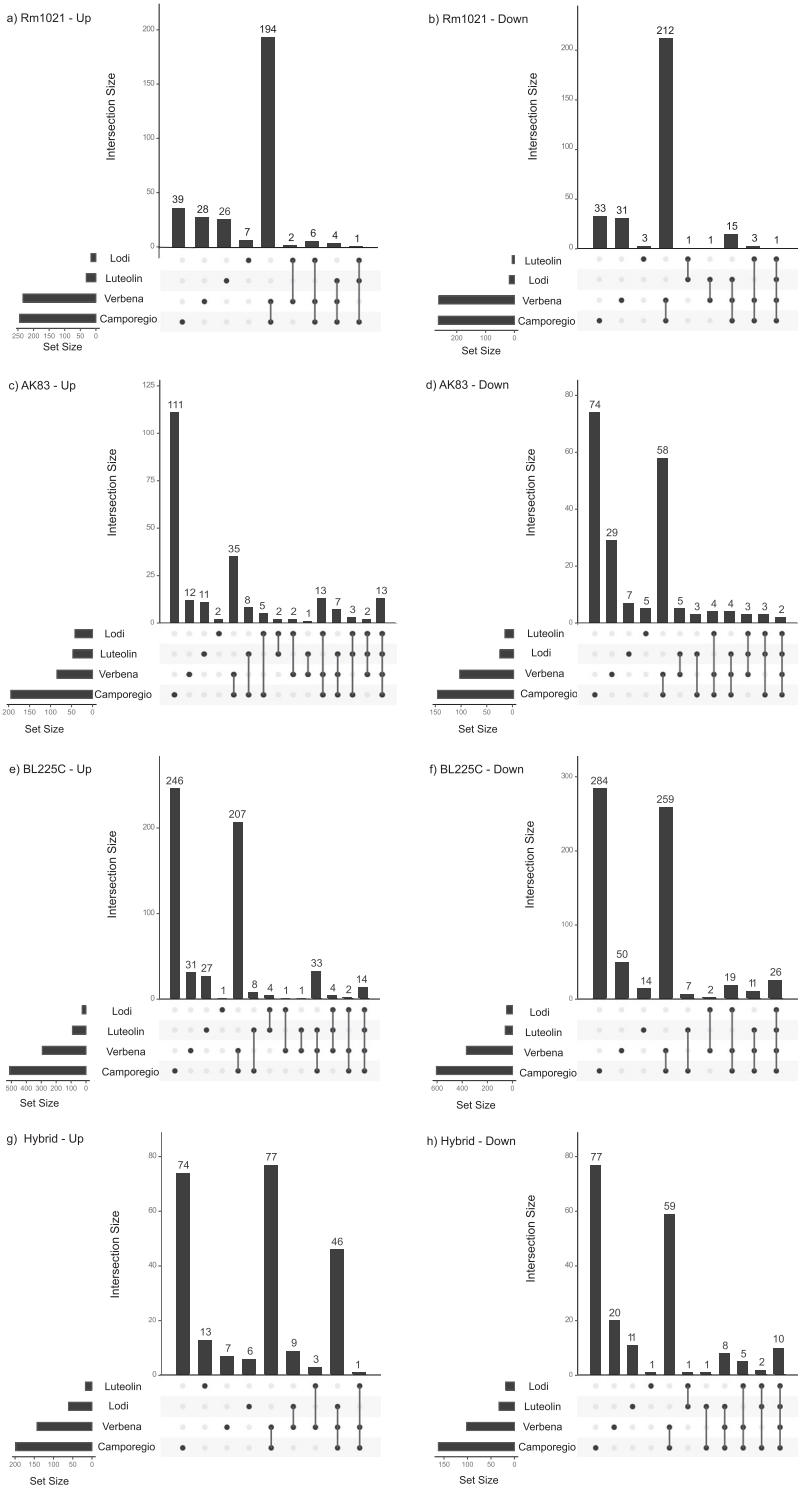


FIG 4 Intersection between up- and downregulated genes under all conditions. The plot reports the numbers of upregulated (a, c, and e) and downregulated (b, d, and f) genes under each condition for each strain. (a and b) Rm1021; (c and d) BL225; (e and f) AK83; (g and h) hybrid strain. Each row of the matrix corresponds to a condition, with the size (number of upregulated/downregulated genes) reported to the left as a bar plot. Each column corresponds to one intersection (similar to a Venn diagram): cells are either empty, indicating that up- or downregulated genes under the specified conditions are not part of the intersection, or filled, indicating that the genes present under the specified conditions are participating at the intersection. Bars on the top show the size of the intersection reported on the bottom.

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

TABLE 2 Number of expressed genes that showed statistical evidence of each type of expression pattern[a]

Parameter	Value for groups of DEGs with significant association								
	Strain	Condition	Strain and condition	Strain × condition[b]				None	Total
Model effect									
Strain	*		*	*	*				
Condition		*	*	*		*			
Strain × condition				*	*	*	*		
No. of DEGs (%)	2,028 (25)	87 (1)	1,201 (15)	1,807 A	436 B	34 C	24 D	2,417 (30)	8,034 (100)

[a]Differential expression related to strain, condition, both strain and condition, or the interaction between strain and condition is reported. Percentages are calculated based on the total number of DEGs. Significance was based on a false discovery rate (FDR)-corrected *P* value of <0.05. * indicates whether the effect of the tested model on the expression of the gene is significant. A gene can be associated with strain (gene differentially expressed only between strains), condition (gene differentially expressed only between different conditions), strain and condition only (gene differentially expressed in relation to strain and condition but not considering the full model strain × condition), or the interaction between strain and condition (strain × condition column). The last situation can be due to a significant association with the three tested models (A), the full model and one of the others (B and C), or the full model only (D).

[b]The total number of DEGs found for strain and condition was 2,301 (29%).

succinoglucan biosynthesis genes such as *wgaA* (*sm_b21319*) and *wgeA* (*sm_b21314*). The Verbena and Camporegio root exudates induced the expression of the rhizobactin transport gene (*sma2337* [*rhtX*]) of Rm1021 and BL225C; this gene is not found in AK83. This may be a consequence of the root exudates chelating the available iron (55), consequently eliciting siderophore production that can inhibit the growth of strains lacking siderophores (56). Plasmid pSINME01 of *S. meliloti* AK83 exhibits similarity with plasmid pHRC017 of *S. meliloti* C017, which confers a competitive advantage for nodule occupancy and host range restrictions (57). Considering that a few of the genes on the plasmids pSINME01 and pSINME02 were differentially expressed upon exposure to root exudates, it is possible that the accessory plasmids of strain AK83 also contribute to competition for rhizosphere colonization (57).

Differences in gene expression patterns across conditions may be related, in part, to differences in the presence of flavonoids. In *S. meliloti*, it is known that root exudates containing flavone molecules activate the transcriptional regulator NodD (43), which triggers the synthesis of Nod factor required for nodule formation. To gain insight into the influence of NodD on the observed stimulons, we compared the *S. meliloti* Rm1021 data to those of the well-known regulons of NodD1 (requiring plant compounds for its activation) and NodD3 (not requiring plant compounds but relying on indirect activation through SyrM and NodD1 [58]) established previously (51, 59). We found that out of the 26 genes of the NodD1 regulon, 7 and 6 were observed in the DEGs in response to the Verbena and Camporegio root exudates, respectively. Camporegio and Verbena root extracts putatively contained apigenin, while the Lodi root extract lacked apigenin, suggesting a role of apigenin in the differential expression pattern observed. For the 226 genes of the NodD3 regulon, 105, 104, and 4 were found in the DEGs in response to the Verbena, Camporegio, and Lodi root exudates, respectively. The presence of a partial overlap of the known *nod* regulons (~20% or fewer of the DEGs under each condition) suggests that most of the observed DEGs belong to *nod*-independent regulons. Moreover, some of these genes showed contrasting patterns of expression, suggesting that the root exudates may also contain antagonistic molecules that repress the *nod* regulon, as previously reported (43, 60).

Mobilization of the symbiotic megaplasmid results in nonadditive changes in stimulons. To evaluate the impact of interreplicon epistatic interactions on the transcriptional response of *S. meliloti* to alfalfa root exudates, we used RNA-seq to characterize the response of a previously constructed *S. meliloti* hybrid strain containing the symbiotic megaplasmid (pSINMEB01) of strain BL225C (48). Cluster analyses clearly demonstrated that the transcriptome (both global and restricted to pSymA-pSINMEB01 orthologs only) of the hybrid strain differed from those of both the BL225C and Rm1021 wild-type strains under all conditions (Fig. 3b; Fig. S4). Of particular interest were the results observed during exposure to the Lodi root exudate. We previously showed that alfalfa cv. Lodi plants inoculated with the hybrid strain were larger than those inoculated

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

TABLE 3 Selected COG categories over- or underrepresented among the DEGs[a]

	Log_2 fold change											
	Luteolin			Camporegio			Verbena			Lodi		
COG category	Rm1021	BL225C	AK83	Rm1021	BL225C	AK83	Rm1021	BL225C	AK83	Rm1021	BL225C	AK83
Upregulated genes												
G	—	—	—	−1.14	−1.59	−1.67	−1.30	−1.64	—	—	—	—
J	—	—	—	2.81	0.92	2.50	2.80	1.16	1.81	—	—	—
N	—	—	—	4.19	—	—	4.25	—	—	—	—	—
O	—	—	2.38	1.56	1.44	1.50	1.44	1.91	2.21	—	—	2.96
Downregulated genes												
C	—	—	—	1.88	0.88	1.74	1.89	1.11	1.79	2.71	—	—

[a]Values represent the log_2 fold changes in the abundances of genes annotated with the given COG category relative to the amount expected by chance. Dashes indicate that the COG category is not statistically different than chance under the given condition (significance threshold of a *P* value of ≤0.05).

with either BL225C or Rm2011 (48). Here, we observed that exposure to the Lodi root exudate results in more differentially expressed genes in the hybrid strain (98 genes) than in either Rm1021 or BL225C (32 and 76 genes, respectively) (Table 1). In particular, a cluster of genes was specifically upregulated in the hybrid strain, and the majority of these genes were located on the symbiotic megaplasmid. This peculiar feature of the Lodi-induced transcriptome in the hybrid strain was also highlighted by the cluster analysis of pSymA-pSINMEB01 orthologs; only in the hybrid strain did the Lodi-induced expression profile cluster with those of the Camporegio and Verbena root exudates (Fig. 3b). The presence of these (possibly nonadditive) transcriptional changes may reflect a loss of *cis*-regulation of these megaplasmid genes by chromosomal regulators (61, 62), providing a potential molecular mechanism underlying the improved symbiotic phenotype of the hybrid compared to both wild-type strains.

DISCUSSION

Rhizobium-legume interactions are complex multistep phenomena that begin with an exchange of signals between two partners (18, 63). The rhizobia initially detect the plant through the perception of flavonoids in the root exudate of legumes by NodD proteins, which then triggers the production of lipochitooligosaccharide molecules known as Nod factors. Nod factors are then recognized by specific LysM receptor kinase proteins in plant root cells, triggering the symbiosis signaling pathway and initiating the formation of a nodule. However, root exudates contain a mixture of flavonoids, some of them having different agonistic activities on NodD (43). Root exudates also contain many other molecules that can serve as signals or support rhizobium metabolism, such as amino acids and sugars, that may influence the ability of rhizobia to successfully colonize the rhizosphere and be in a position to enter the symbiosis (64, 65). Consequently, interactions between plant and rhizobium genotypes are expected to influence the success of the initial interaction between the two partners.

Previous works have identified a clear role for GxG interactions in the partnership between *S. meliloti* and *Medicago truncatula* (66), demonstrating that aerial biomass was influenced by the plant and rhizobium genotypes as well as their interaction. Here, we demonstrated that GxG interactions also have a significant impact on the adherence of *S. meliloti* strains to alfalfa roots, as a representative phenotype for an early stage of the interaction between these partners. Rhizosphere colonization appears to have a direct impact on nodule colonization (49, 67); while our data do not address if root adhesion is correlated with competition for nodule occupancy in mixed inocula, they suggest that root adhesion is poorly correlated with overall symbiotic efficiency in single-inoculum studies. Previous studies have demonstrated the influence of GxG interactions on the nodule transcriptome of *Medicago-Sinorhizobium* symbioses (35, 36). Here, we showed that GxG interactions similarly have an important contribution in determining the transcriptional response of *S. meliloti* to the detection of *Medicago sativa* root exudates.

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

Together, these results demonstrate that GxG interactions have a meaningful impact on the outcome of rhizobium-legume symbioses at multiple stages of development.

The exposure of *B. diazoefficiens* to soybean root exudates resulted in changes in the expression of 450 genes, representing nearly 5.6% of the genome, and the impacts of soybean root exudates differed between the two tested *B. diazoefficiens* strains (53). Similarly, between 0.5% and 20% of *S. meliloti* genes were differentially expressed following exposure to alfalfa root exudates, depending on the host-symbiont combination. The similarities/differences in the responses of the three *S. meliloti* strains to treatments did not appear to depend on the phylogenetic relatedness of the strains (68), although this cannot be definitively concluded without analysis of additional strains. Nevertheless, these results emphasize the importance of transcriptional rewiring during strain diversification in bacteria (62). Similarly, studies with eukaryotic organisms indicate that adaptation has an important role in differentiating the gene expression patterns of organisms (69, 70).

The root exudate stimulons only partially overlapped the stimulons of luteolin, a known inducer of NodD in *S. meliloti* (43), confirming that alfalfa root exudates contain numerous molecular signals aside from flavonoids that may influence the competitiveness of various rhizobium strains. Importantly, the transcriptional patterns induced by alfalfa root exudates differed depending on the cultivar from which they were collected; whether these differences are adaptive requires further investigation. Additionally, although root exudate metabolomic analysis was mainly descriptive, and relatively few peaks could be identified, there was a similar pattern between the differences in the *S. meliloti* gene expression profiles and the overall chemical similarity of the root exudates as measured by LC-MS; the Camporegio and Verbena root exudates induced similar gene expression changes while also being similar along the second principal component of variance (accounting for 30% of the variance) in the PCA of the root exudate composition. In future work, it would be interesting to define which compounds in the root exudates have the greatest impact on the *S. meliloti* transcriptome.

In addition to the impact of GxG interactions on rhizobium-legume symbioses, there is the potential for interreplicon interactions within rhizobium genomes to further influence the symbiosis. Indeed, interreplicon epistatic interactions are abundant in the *S. meliloti* genome (71). To address the contribution of interreplicon interactions to symbiosis, we examined a hybrid strain in which the symbiotic megaplasmid of *S. meliloti* Rm2011 (a strain nearly identical to Rm1021 [72]) was replaced with the symbiotic megaplasmid of *S. meliloti* BL225C. Nonadditive effects on the transcriptional profiles associated with all three replicons were observed in the hybrid strain relative to Rm1021 and BL225C, indicating that megaplasmid mobilization induced a global rewiring of gene expression, likely due to transcriptional cross talk among the replicons (62, 73). Similarly, nonadditive effects on the transcriptome of plant hybrids have been extensively explored (74) and demonstrated as one of the bases for heterosis in crops (75). In previous work looking for regulatory modules where the transcription factor and target genes reside on different replicons in *S. meliloti* Rm1021 (62), we found 17 transcriptional regulators encoded by the chromosome or chromid with predicted target genes on the megaplasmid. Among those transcription factors, systems related to exopolysaccharide production (ExpG), transport (PcaQ), and metabolism (IolR and GlnBK) were present, supporting the hypothesis of a global rewiring of gene expression networks and a wide range of effects of this rewiring. The results with the hybrid led us to hypothesize that the large symbiotic variability observed in natural *S. meliloti* isolates may partly be related to genome-wide transcriptome changes following large-scale horizontal gene transfer followed by natural selection. Moreover, we speculate that while the megaplasmid is the key element for a general response (i.e., cultivar independent) to species-specific host plant associations, the rhizobium chromosome and chromid fine-tune these responses in a genotype-dependent manner. If true, however, this would limit our ability to predict the competitiveness of rhizobium isolates from their simple genome sequence; instead, a more complex understanding of global regulatory network control would be required.

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

In conclusion, this study demonstrated that the initial perception of legumes by rhizobia leads to hundreds of changes in the rhizobium transcriptome and that these changes are dependent on the plant genotype, the rhizobium genotype, and genotype-by-genotype interactions. These results complement previous studies demonstrating the role of GxG interactions in determining the transcriptome of both the legume and rhizobium partners in mature N_2-fixing nodules (35, 36). The majority of genes upregulated in response to alfalfa root exudates were conserved in all three strains, supporting the hypothesis that the *S. meliloti* lineage was adapted to rhizosphere colonization before gaining the genes required for symbiotic nitrogen fixation (49). Additionally, the transcriptional response to the perception of alfalfa root exudates involved genes from all three of the *S. meliloti* replicons and seemingly involved nonadditive effects resulting from interreplicon interactions.

MATERIALS AND METHODS

Microbiological methods and plant assays. A list of strains, their host plants of origin, as well as the plant varieties used is reported in Table S4 in the supplemental material. Plant varieties included three contrasting alfalfa genotypes: *Medicago falcata* (Verbena), *M. sativa* (Lodi), and *Medicago* × *varia* (*M. sativa* × *M. falcata*). Strains included *S. meliloti* Rm1021, BL225C, AK83, and a hybrid strain containing the chromosome and pSymB of strain Rm2011 and the symbiotic megaplasmid (pSINMEB01) of strain BL225C. *S. meliloti* Rm2011 is nearly isogenic to Rm1021, both being independent streptomycin-resistant derivatives of the nodule isolate SU47 (76, 77). Details on strains, plant growth, and symbiotic assays are found in Text S1 in the supplemental material. The root adhesion test was performed 5 days following the inoculation of plantlets (Text S1). Differences were evaluated by one-way analysis of variance (ANOVA) Tukey pairwise contrast and using the Scott-Knott procedure as implemented in R (78). All primer pairs used are reported in Table S4.

Root exudate production and metabolomic analyses. Root exudates were produced by growing plants under sterile conditions in water for 14 days, as previously reported (79) and as reported in Text S1. Elemental analysis (CHNS) was performed on crude root exudates (a combined sample for each cultivar) using a carbon, hydrogen, and nitrogen analyzer (CHN-S Flash E1112; Thermo Finnigan, San Jose, CA, USA). Metabolomic analysis was performed by LC-MS, and data from reverse-phase UPLC (RP-UPLC) and UPLC-MS were combined to build the final data matrix. Principal-component analysis (PCA) was performed on the Bray-Curtis dissimilarity obtained from each peak identification (ID) value (Text S1). Statistical differences in single metabolites were assessed by Simper analysis based on the decomposition of the Bray-Curtis dissimilarity obtained from each peak ID value. All statistical analyses were done with the vegan package of R (80). The PubChem database was used for additional peak identification from brute formulas (https://pubchem.ncbi.nlm.nih.gov/).

RNA isolation and RNA sequencing. Cultures of *S. meliloti*, grown overnight in M9-succinate medium at 30°C at 130 rpm, were diluted to an optical density at 600 nm (OD_{600}) of 0.05 in 5 ml of M9-succinate medium and incubated until an OD_{600} of 0.4 was reached. Next, either 10 μM luteolin (Sigma-Aldrich) or one of the alfalfa root exudates (normalized by the total organic carbon as measured by the CHNS analysis) was added to each of the cultures, and the mixture was incubated for an additional 4 h at 30°C with shaking at 130 rpm. Biological replicates were performed for each of the three strains across the five conditions. Total RNA was extracted using RNeasy minikits (Qiagen) from 0.5 ml of the culture and subjected to DNase I treatment. Details on the RNA isolation procedure and quality checks are provided in Text S1. Validation of expression differences was done using reverse transcriptase qPCR as described in Text S1. Protocols for rRNA depletion and library construction are described in Text S1. Libraries were sequenced on an Illumina Novaseq 6000 apparatus with an SP flow cell.

Read mapping, counting, and differential expression analysis. Trimmed and demultiplexed reads were mapped back to transcripts using Salmon (version 1.1.0) (81) (see Text S1 for details). Quantification files produced by Salmon were then imported into R using the tximport package (version 1.10.1) (82). Differential abundance analysis was performed with the DESeq2 version 1.22.2 package (83) on single strains under different conditions.

Statistical analysis of differentially expressed genes. For each *S. meliloti* strain, genes differentially expressed ($\log_2$ fold change of $\geq$1; P value of $<$0.01) under at least one condition relative to the control conditions were identified, and all fold change values for these genes were extracted. To compare expression values of genes conserved between Rm1021, AK83, and BL225C, the pangenome of the three strains was calculated using Roary version 3.13.0 (84) with an identity threshold of 90%, and the genes found in all three strains (the core genes) were recorded. Under each condition, core genes differentially expressed in at least one strain relative to the control conditions were identified, and the fold change values for the gene and its orthologs in the other strains were extracted.

All genes of *S. meliloti* strains Rm1021, AK83, and BL225C were functionally annotated using standalone version 2 of eggNOG-mapper (85, 86) with default settings and the following two modifications: the mode was set to diamond, and query cover was set to 20. Methods for the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Clusters of Orthologous Genes (COG) category annotations are reported in Text S1.

Nested likelihood ratio tests (LRTs) were used to evaluate the statistical significance of strain, condition, and strain-by-condition interaction effects on gene expression. Transcripts were collapsed into

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

orthologous groups based on the output of Roary, as described above. Counts produced by Salmon were collapsed following the group ID provided by Roary, producing a single table with ortholog-level quantification of transcripts. The produced table was then used to perform a nested LRT with DESeq2. Strains and conditions were used together with their interaction to build a model for each group. Terms were then removed one by one to test their impact on the likelihood of the full model (as described in the DESeq2 documentation at http://bioconductor.org/packages/devel/bioc/vignettes/DESeq2/inst/doc/DESeq2.html#likelihood-ratio-test).

Data availability. Gene expression data are available at GEO under the accession number GSE151705. Custom scripts developed for this work can be found in the GitHub repository at https://github.com/hyhy8181994/Sinorhizobium-RNAseq-2020.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

TEXT S1, DOCX file, 0.04 MB.
FIG S1, PDF file, 0.04 MB.
FIG S2, PDF file, 0.2 MB.
FIG S3, PDF file, 0.1 MB.
FIG S4, PDF file, 0.1 MB.
TABLE S1, PDF file, 0.02 MB.
TABLE S2, PDF file, 0.01 MB.
TABLE S3, CSV file, 0.01 MB.
TABLE S4, PDF file, 0.1 MB.
DATA SET S1, XLSX file, 0.8 MB.

ACKNOWLEDGMENTS

We are grateful to Gabriele Brazzini for technical assistance in setting up the root adhesion test and Susanna Pucci, Centro di Microanalisi, University of Florence, Italy, for CHNS analysis.

This work was supported by Fondazione Cassa di Risparmio di Firenze grant number 18204, 2017.0719; by a MICRO4Legumes grant (Italian Ministry of Agriculture); and by a grant for the Dipartimento di Eccellenza 2018-2022 by the Italian Ministry of Education, University, and Research (MIUR). L.C. was supported by a fellowship from MICRO4Legumes (Italian Ministry of Agriculture). Work in the G.C.D. laboratory is supported by funding from Queen's University and the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. Rosenberg E, Zilber-Rosenberg I. 2016. Microbes drive evolution of animals and plants: the hologenome concept. mBio 7:e01395-15. https://doi.org/10.1128/mBio.01395-15.
2. Mus F, Crook MB, Garcia K, Costas AG, Geddes BA, Kouri ED, Paramasivan P, Ryu M-H, Oldroyd GED, Poole PS, Udvardi MK, Voigt CA, Ané J-M, Peters JW. 2016. Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. Appl Environ Microbiol 82:3698–3710. https://doi.org/10.1128/AEM.01055-16.
3. Lee YK, Mazmanian SK. 2010. Has the microbiota played a critical role in the evolution of the adaptive immune system? Science 330:1768–1773. https://doi.org/10.1126/science.1195568.
4. Jones BW, Nishiguchi MK. 2004. Counterillumination in the Hawaiian bobtail squid, Euprymna scolopes Berry (Mollusca: Cephalopoda). Mar Biol 144:1151–1155. https://doi.org/10.1007/s00227-003-1285-3.
5. Nyholm SV, McFall-Ngai M. 2004. The winnowing: establishing the squid-vibrio symbiosis. Nat Rev Microbiol 2:632–642. https://doi.org/10.1038/nrmicro957.
6. Kiers ET, Rousseau RA, West SA, Denison RF. 2003. Host sanctions and the legume-rhizobium mutualism. Nature 425:78–81. https://doi.org/10.1038/nature01931.
7. Mwafulirwa L, Baggs EM, Russell J, George T, Morley N, Sim A, de la Fuente Cantó C, Paterson E. 2016. Barley genotype influences stabilization of rhizodeposition-derived C and soil organic matter mineralization. Soil Biol Biochem 95:60–69. https://doi.org/10.1016/j.soilbio.2015.12.011.
8. Escudero-Martinez C, Bulgarelli D. 2019. Tracing the evolutionary routes of plant-microbiota interactions. Curr Opin Microbiol 49:34–40. https://doi.org/10.1016/j.mib.2019.09.013.
9. Pasolli E, Asnicar F, Manara S, Zolfo M, Karcher N, Armanini F, Beghini F, Manghi P, Tett A, Ghensi P, Collado MC, Rice BL, DuLong C, Morgan XC, Golden CD, Quince C, Huttenhower C, Segata N. 2019. Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. Cell 176:649–662.e20. https://doi.org/10.1016/j.cell.2019.01.001.
10. Singh BK, Liu H, Trivedi P. 2020. Eco-holobiont: a new concept to identify drivers of host-associated microorganisms. Environ Microbiol 22:564–567. https://doi.org/10.1111/1462-2920.14900.
11. de Souza RSC, Armanhi JSL, Damasceno NDB, Imperial J, Arruda P. 2019. Genome sequences of a plant beneficial synthetic bacterial community reveal genetic features for successful plant colonization. Front Microbiol 10:1779. https://doi.org/10.3389/fmicb.2019.01779.
12. Levy A, Gonzalez IS, Mittelviefhaus M, Clingenpeel S, Paredes SH, Miao J, Wang K, Devescovi G, Stillman K, Monteiro F, Alvarez BR, Lundberg DS, Lu T-Y, Lebeis S, Jin Z, McDonald M, Klein AP, Feltcher ME, Rio TG, Grant SR, Doty SL, Ley RE, Zhao B, Venturi V, Pelletier DA, Vorholt JA, Tringe SG, Woyke T, Dangl JL. 2018. Genomic features of bacterial adaptation to plants. Nat Genet 50:138–150. https://doi.org/10.1038/s41588-017-0012-9.
13. Harrison E, Brockhurst MA. 2012. Plasmid-mediated horizontal gene transfer is a coevolutionary process. Trends Microbiol 20:262–267. https://doi.org/10.1016/j.tim.2012.04.003.
14. Batstone RT, O'Brien AM, Harrison TL, Frederickson ME. 2020. Experimental evolution makes microbes more cooperative with their local host genotype. Science 370:476–478. https://doi.org/10.1126/science.abb7222.
15. Pini F, Galardini M, Bazzicalupo M, Mengoni A. 2011. Plant-bacteria

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

association and symbiosis: are there common genomic traits in Alphaproteobacteria? Genes (Basel) 2:1017–1032. https://doi.org/10.3390/genes2041017.
16. Guttman D, McHardy AC, Schulze-Lefert P. 2014. Microbial genome-enabled insights into plant-microorganism interactions. Nat Rev Genet 15:797–813. https://doi.org/10.1038/nrg3748.
17. Poole P, Ramachandran V, Terpolilli J. 2018. Rhizobia: from saprophytes to endosymbionts. Nat Rev Microbiol 16:291–303. https://doi.org/10.1038/nrmicro.2017.171.
18. Oldroyd GED. 2013. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol 11:252–263. https://doi.org/10.1038/nrmicro2990.
19. Gage DJ. 2004. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiol Mol Biol Rev 68:280–300. https://doi.org/10.1128/MMBR.68.2.280-300.2004.
20. Udvardi M, Poole PS. 2013. Transport and metabolism in legume-rhizobia symbioses. Annu Rev Plant Biol 64:781–805. https://doi.org/10.1146/annurev-arplant-050312-120235.
21. Kereszt A, Mergaert P, Kondorosi E. 2011. Bacteroid development in legume nodules: evolution of mutual benefit or of sacrificial victims? Mol Plant Microbe Interact 24:1300–1309. https://doi.org/10.1094/MPMI-06-11-0152.
22. Benezech C, Doudement M, Gourion B. 2020. Legumes tolerance to rhizobia is not always observed and not always deserved. Cell Microbiol 22:e13124. https://doi.org/10.1111/cmi.13124.
23. Gibson KE, Kobayashi H, Walker GC. 2008. Molecular determinants of a symbiotic chronic infection. Annu Rev Genet 42:413–441. https://doi.org/10.1146/annurev.genet.42.110807.091427.
24. Burghardt LT. 2020. Evolving together, evolving apart: measuring the fitness of rhizobial bacteria in and out of symbiosis with leguminous plants. New Phytol 228:28–34. https://doi.org/10.1111/nph.16045.
25. Paffetti D, Scotti C, Gnocchi S, Fancelli S, Bazzicalupo M. 1996. Genetic diversity of an Italian Rhizobium meliloti population from different Medicago sativa varieties. Appl Environ Microbiol 62:2279–2285. https://doi.org/10.1128/AEM.62.7.2279-2285.1996.
26. Carelli M, Gnocchi S, Fancelli S, Mengoni A, Paffetti D, Scotti C, Bazzicalupo M. 2000. Genetic diversity and dynamics of Sinorhizobium meliloti populations nodulating different alfalfa cultivars in Italian soils. Appl Environ Microbiol 66:4785–4789. https://doi.org/10.1128/aem.66.11.4785-4789.2000.
27. Rangin C, Brunel B, Cleyet-Marel J-C, Perrineau M-M, Béna G. 2008. Effects of Medicago truncatula genetic diversity, rhizobial competition, and strain effectiveness on the diversity of a natural Sinorhizobium species community. Appl Environ Microbiol 74:5653–5661. https://doi.org/10.1128/AEM.01107-08.
28. Heath KD, Tiffin P. 2007. Context dependence in the coevolution of plant and rhizobial mutualists. Proc Biol Sci 274:1905–1912. https://doi.org/10.1098/rspb.2007.0495.
29. Heath KD, Tiffin P. 2009. Stabilizing mechanisms in a legume-rhizobium mutualism. Evolution 63:652–662. https://doi.org/10.1111/j.1558-5646.2008.00582.x.
30. Heath KD. 2010. Intergenomic epistasis and coevolutionary constraint in plants and rhizobia. Evolution 64:1446–1458. https://doi.org/10.1111/j.1558-5646.2009.00913.x.
31. Ehinger M, Mohr TJ, Starcevich JB, Sachs JL, Porter SS, Simms EL. 2014. Specialization-generalization trade-off in a Bradyrhizobium symbiosis with wild legume hosts. BMC Ecol 14:8. https://doi.org/10.1186/1472-6785-14-8.
32. Burghardt LT, Epstein B, Guhlin J, Nelson MS, Taylor MR, Young ND, Sadowsky MJ, Tiffin P. 2018. Select and resequence reveals relative fitness of bacteria in symbiotic and free-living environments. Proc Natl Acad Sci U S A 115:2425–2430. https://doi.org/10.1073/pnas.1714246115.
33. Burghardt LT, Trujillo DI, Epstein B, Tiffin P, Young ND. 2020. A select and resequence approach reveals strain-specific effects of Medicago nodule-specific PLAT-domain genes. Plant Physiol 182:463–471. https://doi.org/10.1104/pp.19.00831.
34. Roux B, Rodde N, Jardinaud MF, Timmers T, Sauviac L, Cottret L, Carrère S, Sallet E, Courcelle E, Moreau S, Debellé F, Capela D, De Carvalho-Niebel F, Gouzy J, Bruand C, Gamas P. 2014. An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. Plant J 77:817–837. https://doi.org/10.1111/tpj.12442.
35. Heath KD, Burke PV, Stinchcombe JR. 2012. Coevolutionary genetic variation in the legume-rhizobium transcriptome. Mol Ecol 21:4735–4747. https://doi.org/10.1111/j.1365-294X.2012.05629.x.
36. Burghardt LT, Guhlin J, Chun CL, Liu J, Sadowsky MJ, Stupar RM, Young ND, Tiffin P. 2017. Transcriptomic basis of genome by genome variation in a legume-rhizobia mutualism. Mol Ecol 26:6122–6135. https://doi.org/10.1111/mec.14285.
37. Triplett EW, Sadowsky MJ. 1992. Genetics of competition for nodulation of legumes. Annu Rev Microbiol 46:399–422. https://doi.org/10.1146/annurev.mi.46.100192.002151.
38. Checcucci A, DiCenzo GC, Bazzicalupo M, Mengoni A. 2017. Trade, diplomacy, and warfare: the quest for elite rhizobia inoculant strains. Front Microbiol 8:2207. https://doi.org/10.3389/fmicb.2017.02207.
39. Sprent JI, Ardley J, James EK. 2017. Biogeography of nodulated legumes and their nitrogen-fixing symbionts. New Phytol 215:40–56. https://doi.org/10.1111/nph.14474.
40. Frame J, Charlton JFL, Laidlaw AS. 1998. Temperate forage legumes. CAB International, Wallingford, United Kingdom.
41. Galibert F, Finan TM, Long SR, Puhler A, Abola P, Ampe F, Barloy-Hubler F, Barnett MJ, Becker A, Boistard P, Bothe G, Boutry M, Bowser L, Buhrmester J, Cadieu E, Capela D, Chain P, Cowie A, Davis RW, Dreano S, Federspiel NA, Fisher RF, Gloux S, Godrie T, Goffeau A, Golding B, Gouzy J, Gurjal M, Hernandez-Lucas I, Hong A, Huizar L, Hyman RW, Jones T, Kahn D, Kahn ML, Kalman S, Keating DH, Kiss E, Komp C, Lelaure V, Masuy D, Palm C, Peck MC, Pohl TM, Portetelle D, Purnelle B, Ramsperger U, Surzycki R, Thebault P, Vandenbol M, Vorholter FJ, et al. 2001. The composite genome of the legume symbiont Sinorhizobium meliloti. Science 293:668–672. https://doi.org/10.1126/science.1060966.
42. Harrison PW, Lower RPJ, Kim NKD, Young JPW. 2010. Introducing the bacterial "chromid": not a chromosome, not a plasmid. Trends Microbiol 18:141–148. https://doi.org/10.1016/j.tim.2009.12.010.
43. Peck MC, Fisher RF, Long SR. 2006. Diverse flavonoids stimulate NodD1 binding to nod gene promoters in Sinorhizobium meliloti. J Bacteriol 188:5417–5427. https://doi.org/10.1128/JB.00376-06.
44. Galardini M, Mengoni A, Brilli M, Pini F, Fioravanti A, Lucas S, Lapidus A, Cheng J-F, Goodwin L, Pitluck S, Land M, Hauser L, Woyke T, Mikhailova N, Ivanova N, Daligault H, Bruce D, Detter C, Tapia R, Han C, Teshima H, Mocali S, Bazzicalupo M, Biondi EG. 2011. Exploring the symbiotic pangenome of the nitrogen-fixing bacterium Sinorhizobium meliloti. BMC Genomics 12:235. https://doi.org/10.1186/1471-2164-12-235.
45. Biondi EG, Tatti E, Comparini D, Giuntini E, Mocali S, Giovannetti L, Bazzicalupo M, Mengoni A, Viti C. 2009. Metabolic capacity of Sinorhizobium (Ensifer) meliloti strains as determined by phenotype microarray analysis. Appl Environ Microbiol 75:5396–5404. https://doi.org/10.1128/AEM.00196-09.
46. Checcucci A, Azzarello E, Bazzicalupo M, Galardini M, Lagomarsino A, Mancuso S, Marti L, Marzano MC, Mocali S, Squartini A, Zanardo M, Mengoni A. 2016. Mixed nodule infection in Sinorhizobium meliloti-Medicago sativa symbiosis suggest the presence of cheating behavior. Front Plant Sci 7:835. https://doi.org/10.3389/fpls.2016.00835.
47. Caetano-Anollés G, Crist-Estes DK, Bauer WD. 1988. Chemotaxis of Rhizobium meliloti to the plant flavone luteolin requires functional nodulation genes. J Bacteriol 170:3164–3169. https://doi.org/10.1128/jb.170.7.3164-3169.1988.
48. Checcucci A, DiCenzo GC, Ghini V, Bazzicalupo M, Becker A, Decorosi F, Döhlemann J, Fagorzi C, Finan TM, Fondi M, Luchinat C, Turano P, Vignolini T, Viti C, Mengoni A. 2018. Creation and characterization of a genomically hybrid strain in the nitrogen-fixing symbiotic bacterium Sinorhizobium meliloti. ACS Synth Biol 7:2365–2378. https://doi.org/10.1021/acssynbio.8b00158.
49. Salas ME, Lozano MJ, López JL, Draghi WO, Serrania J, Torres Tejerizo GA, Albicoro FJ, Nilsson JF, Pistorio M, Del Papa MF, Parisi G, Becker A, Lagares A. 2017. Specificity traits consistent with legume-rhizobia coevolution displayed by Ensifer meliloti rhizosphere colonization. Environ Microbiol 19:3423–3438. https://doi.org/10.1111/1462-2920.13820.
50. Spini G, Decorosi F, Cerboneschi M, Tegli S, Mengoni A, Viti C, Giovannetti L. 2016. Effect of the plant flavonoid luteolin on Ensifer meliloti 3001 phenotypic responses. Plant Soil 399:159–178. https://doi.org/10.1007/s11104-015-2659-2.
51. Barnett MJ, Long SR. 2015. The Sinorhizobium meliloti SyrM regulon: effects on global gene expression are mediated by syrA and nodD3. J Bacteriol 197:1792–1806. https://doi.org/10.1128/JB.02626-14.
52. Webb BA, Helm RF, Scharf BE. 2016. Contribution of individual chemoreceptors to Sinorhizobium meliloti chemotaxis towards amino acids of

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

host and nonhost seed exudates. Mol Plant Microbe Interact 29:231–239. https://doi.org/10.1094/MPMI-12-15-0264-R.
53. Liu Y, Jiang X, Guan D, Zhou W, Ma M, Zhao B, Cao F, Li L, Li J. 2017. Transcriptional analysis of genes involved in competitive nodulation in Bradyrhizobium diazoefficiens at the presence of soybean root exudates. Sci Rep 7:10946. https://doi.org/10.1038/s41598-017-11372-0.
54. McIntosh M, Krol E, Becker A. 2008. Competitive and cooperative effects in quorum-sensing-regulated galactoglucan biosynthesis in Sinorhizobium meliloti. J Bacteriol 190:5308–5317. https://doi.org/10.1128/JB.00063-08.
55. Parker DR, Reichman SM, Crowley DE. 2005. Metal chelation in the rhizosphere, p 57–93. *In* Zobel RW, Wright SF (ed), Roots and soil management: interactions between roots and the soil, vol 48. American Society of Agronomy, Inc, Madison, WI.
56. diCenzo GC, MacLean AM, Milunovic B, Golding GB, Finan TM. 2014. Examination of prokaryotic multipartite genome evolution through experimental genome reduction. PLoS Genet 10:e1004742. https://doi.org/10.1371/journal.pgen.1004742.
57. Crook MB, Lindsay DP, Biggs MB, Bentley JS, Price JC, Clement SC, Clement MJ, Long SR, Griffitts JS. 2012. Rhizobial plasmids that cause impaired symbiotic nitrogen fixation and enhanced host invasion. Mol Plant Microbe Interact 25:1026–1033. https://doi.org/10.1094/MPMI-02-12-0052-R.
58. Maillet F, Debellé F, Dénarié J. 1990. Role of the nodD and syrM genes in the activation of the regulatory gene nodD3, and of the common and host-specific nod genes of Rhizobium meliloti. Mol Microbiol 4:1975–1984. https://doi.org/10.1111/j.1365-2958.1990.tb02047.x.
59. Capela D, Carrere S, Batut J. 2005. Transcriptome-based identification of the Sinorhizobium meliloti NodD1 regulon. Appl Environ Microbiol 71:4910–4913. https://doi.org/10.1128/AEM.71.8.4910-4913.2005.
60. Zuanazzi JAS, Clergeot PH, Quirion J-C, Husson H-P, Kondorosi A, Ratet P. 1998. Production of Sinorhizobium meliloti nod gene activator and repressor flavonoids from Medicago sativa roots. Mol Plant Microbe Interact 11:784–794. https://doi.org/10.1094/MPMI.1998.11.8.784.
61. Ramos JL, Marqués S, Timmis KN. 1997. Transcriptional control of the Pseudomonas TOL plasmid catabolic operons is achieved through an interplay of host factors and plasmid-encoded regulators. Annu Rev Microbiol 51:341–373. https://doi.org/10.1146/annurev.micro.51.1.341.
62. Galardini M, Brilli M, Spini G, Rossi M, Roncaglia B, Bani A, Chiancianesi M, Moretto M, Engelen K, Bacci G, Pini F, Biondi EG, Bazzicalupo M, Mengoni A. 2015. Evolution of intra-specific regulatory networks in a multipartite bacterial genome. PLoS Comput Biol 11:e1004478. https://doi.org/10.1371/journal.pcbi.1004478.
63. Oldroyd GED, Murray JD, Poole PS, Downie JA. 2011. The rules of engagement in the legume-rhizobial symbiosis. Annu Rev Genet 45:119–144. https://doi.org/10.1146/annurev-genet-110410-132549.
64. Barbour WM, Hattermann DR, Stacey G. 1991. Chemotaxis of Bradyrhizobium japonicum to soybean exudates. Appl Environ Microbiol 57:2635–2639. https://doi.org/10.1128/AEM.57.9.2635-2639.1991.
65. Webb BA, Compton KK, Martin JS, Taylor D, Sobrado P, Scharf BE. 2017. Sinorhizobium meliloti chemotaxis to multiple amino acids is mediated by the chemoreceptor McpU. Mol Plant Microbe Interact 30:770–777. https://doi.org/10.1094/MPMI-04-17-0096-R.
66. Mhadhbi H, Jebara M, Limam F, Huguet T, Aouani ME. 2005. Interaction between Medicago truncatula lines and Sinorhizobium meliloti strains for symbiotic efficiency and nodule antioxidant activities. Physiol Plant 124:4–11. https://doi.org/10.1111/j.1399-3054.2005.00489.x.
67. Entcheva P, Phillips DA, Streit WR. 2002. Functional analysis of Sinorhizobium meliloti genes involved in biotin synthesis and transport. Appl Environ Microbiol 68:2843–2848. https://doi.org/10.1128/aem.68.6.2843-2848.2002.
68. Galardini M, Pini F, Bazzicalupo M, Biondi EG, Mengoni A. 2013. Replicon-dependent bacterial genome evolution: the case of Sinorhizobium meliloti. Genome Biol Evol 5:542–558. https://doi.org/10.1093/gbe/evt027.
69. López-Maury L, Marguerat S, Bähler J. 2008. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. Nat Rev Genet 9:583–593. https://doi.org/10.1038/nrg2398.
70. Whitehead A, Crawford DL. 2006. Variation within and among species in gene expression: raw material for evolution. Mol Ecol 15:1197–1211. https://doi.org/10.1111/j.1365-294X.2006.02868.x.
71. diCenzo GC, Benedict AB, Fondi M, Walker GC, Finan TM, Mengoni A, Griffitts JS. 2018. Robustness encoded across essential and accessory replicons of the ecologically versatile bacterium Sinorhizobium meliloti. PLoS Genet 14:e1007357. https://doi.org/10.1371/journal.pgen.1007357.
72. Meade HM, Signer ER. 1977. Genetic mapping of Rhizobium meliloti. Proc Natl Acad Sci U S A 74:2076–2078. https://doi.org/10.1073/pnas.74.5.2076.
73. diCenzo GC, Wellappili D, Golding GB, Finan TM. 2018. Inter-replicon gene flow contributes to transcriptional integration in the *Sinorhizobium meliloti* multipartite genome. G3 (Bethesda) 8:1711–1720. https://doi.org/10.1534/g3.117.300405.
74. Bell GDM, Kane NC, Rieseberg LH, Adams KL. 2013. RNA-seq analysis of allele-specific expression, hybrid effects, and regulatory divergence in hybrids compared with their parents from natural populations. Genome Biol Evol 5:1309–1323. https://doi.org/10.1093/gbe/evt072.
75. Hochholdinger F, Hoecker N. 2007. Towards the molecular basis of heterosis. Trends Plant Sci 12:427–432. https://doi.org/10.1016/j.tplants.2007.08.005.
76. Sallet E, Roux B, Sauviac L, Jardinaud MF, Carrère S, Faraut T, de Carvalho-Niebel F, Gouzy J, Gamas P, Capela D, Bruand C, Schiex T. 2013. Next-generation annotation of prokaryotic genomes with EuGene-P: application to Sinorhizobium meliloti 2011. DNA Res 20:339–353. https://doi.org/10.1093/dnares/dst014.
77. Krol E, Becker A. 2004. Global transcriptional analysis of the phosphate starvation response in Sinorhizobium meliloti strains 1021 and 2011. Mol Genet Genomics 272:1–17. https://doi.org/10.1007/s00438-004-1030-8.
78. R Development Core Team. 2012. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
79. Checcucci A, Azzarello E, Bazzicalupo M, De Carlo A, Emiliani G, Mancuso S, Spini G, Viti C, Mengoni A. 2017. Role and regulation of ACC deaminase gene in Sinorhizobium meliloti: is it a symbiotic, rhizospheric or endophytic gene? Front Genet 8:6. https://doi.org/10.3389/fgene.2017.00006.
80. Oksanen J, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Simpson G, Solymos P, Stevens M, Wagner H. 2013. vegan: community ecology package. R package version 2.0-10.
81. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14:417–419. https://doi.org/10.1038/nmeth.4197.
82. Soneson C, Love MI, Robinson MD. 2015. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res 4:1521. https://doi.org/10.12688/f1000research.7563.1.
83. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8.
84. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693. https://doi.org/10.1093/bioinformatics/btv421.
85. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, von Mering C, Bork P. 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res 47:D309–D314. https://doi.org/10.1093/nar/gky1085.
86. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, Von Mering C, Bork P. 2017. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. Mol Biol Evol 34:2115–2122. https://doi.org/10.1093/molbev/msx148.

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

International Journal of
Molecular Sciences

MDPI

Review

Scent of a Symbiont: The Personalized Genetic Relationships of Rhizobium—Plant Interaction

Lisa Cangioli, Francesca Vaccaro, Margherita Fini, Alessio Mengoni * and Camilla Fagorzi

Department of Biology, University of Florence, Via Madonna del Piano 6, 50019 Sesto Fiorentino, Italy; lisa.cangioli@unifi.it (L.C.); francesca.vaccaro@unifi.it (F.V.); margherita.fini@stud.unifi.it (M.F.); camilla.fagorzi@unifi.it (C.F.)

* Correspondence: alessio.mengoni@unifi.it

Abstract: Many molecular signals are exchanged between rhizobia and host legume plants, some of which are crucial for symbiosis to take place, while others are modifiers of the interaction, which have great importance in the competition with the soil microbiota and in the genotype-specific perception of host plants. Here, we review recent findings on strain-specific and host genotype-specific interactions between rhizobia and legumes, discussing the molecular actors (genes, gene products and metabolites) which play a role in the establishment of symbiosis, and highlighting the need for research including the other components of the soil (micro)biota, which could be crucial in developing rational-based strategies for bioinoculants and synthetic communities' assemblage.

Keywords: rhizobium genomics; symbiotic interaction; nitrogen fixation; sustainable agriculture

check for updates

Citation: Cangioli, L.; Vaccaro, F.; Fini, M.; Mengoni, A.; Fagorzi, C. Scent of a Symbiont: The Personalized Genetic Relationships of Rhizobium—Plant Interaction. *Int. J. Mol. Sci.* **2022**, *23*, 3358. https://doi.org/10.3390/ijms23063358

Academic Editor: Shawn A. Christensen

Received: 25 February 2022
Accepted: 18 March 2022
Published: 20 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Plants, like other multicellular eukaryotes, live in close association with many other organisms, and they often establish mutualistic relationships with them which, in turn, shape their phenotype. Indeed, many host plant traits are now recognized as being influenced by host-associated organisms, especially microorganisms, either prokaryotic and eukaryotic [1]. The plant phenotype is a result of multiple-level interactions, starting with the host genotype, and including physicochemical environmental factors, but also, more importantly, the microbiome composition, behavior and function [2,3]. There is a large debate on how to consider the heritability of microbe-related traits [2]. It has been suggested that phenotypic variation should be considered as the sum of environmental variance and of additive and nonadditive contributions. These contributions derive from the genotype of the host and of the mutualistic microbes (including their mobile genetic elements, which largely vary among strains) and from the intergenomic epistasis between hosts and microbes [4].

Disentangling these multi-layered, interconnected interactions is complicated. Research can be performed on natural populations by applying statistical methods to disentangle the contribution of the variables [5], or by using model systems in controlled conditions [6]. However, the extreme diversity (genetic and functional) of the native host-associated microbiome limits the possibility of understanding the molecular and genetic factors of the interactions, although relevant works addressing the role of plant genotypes and single genes have been conducted [7–10].

The rhizobia–legume plant mutualistic interaction offers a simplified, though still complex, system to address questions on the heritability of microbe-related traits delving deep into deciphering the genetic determinants and the molecular communication in the establishment of successful associations. Rhizobia are ubiquitous soil bacteria, belonging to the Alpha and Beta Proteobacteria classes, which colonize the rhizosphere and endosphere of different plant species and form symbiotic associations with leguminous plants [11]. The symbiotic association typically starts with the perception of plant-derived flavonoids

that bind the bacterial NodD protein, which, in turn, activates the transcription of *nod* genes, resulting in the production of lipochitooligosaccharides (LCOs) termed Nod factors (NFs). NFs are perceived by the host legume and trigger the division of inner plant cortical cells, resulting in the formation of nodule primordia, and root hair curling which induces infection thread formation [12]. This process facilitates the accommodation of rhizobia inside the plant cell, where they differentiate into bacteroids and start fixing atmospheric nitrogen [12]. Metabolic rewiring and the stress response play crucial roles in symbiotic establishment [13,14], since bacteroids must cope with osmotic stress, changes in pH, reduced oxygen tension [14] and plant peptides. Plant peptides are becoming increasingly relevant since they induce and maintain growth inhibition and can modify host compatibility [15–19]. Related to nodule growth and correlated with bacteroid differentiation, nodules can be found in determinate and indeterminate forms. Determinate nodules are found in tropical and subtropical legumes such as *Lotus japonicus*, soybean (*Glycine max*), common bean (*Phaseolus vulgaris*) and cowpea (*Vigna unguiculata*). Temperate legumes tend to form indeterminate, and one of the most studied models is the legume *Medicago truncatula* and its symbiotic rhizobium *Sinorhizobium meliloti*. Indeed, many of the genes required by the plant and the symbiotic bacterium for establishing a successful symbiosis have been discovered in this model, and genetic tools that enable testing hypotheses have been developed [20,21]. Therefore, while this review aims to be general, most of the research data and examples provided are related to indeterminate nodules and to the model system *M. truncatula*/*S. meliloti*. This system offers excellent working models to molecularly understand host–microbe relationships and host genotype–microbial genotype interactions [20].

2. Evidence of Genotypic Signature in the Rhizobium—Plant Partner Choice

Several population genetics investigations on rhizobial symbionts have been performed on many species, spanning from soybean symbionts to fava bean and alfalfa symbionts, initially with PCR-based approaches, then more recently with genome sequencing (see, for instance, [22–30]). Since the first works, it was clear that strains isolated from nodules of different host plant genotypes (e.g., varieties, or closely related species) cluster, to some extent, into different groups. More than 20 years ago, a long-term experiment where physically separated nodulating populations were followed over the years on the same plants showed increasing relationships with the host plant genotypes, though random genetic drift was also an important factor shaping the rhizobial population structure [22]. This long-term evolution experiment with combinations of different cultivars of *Medicago sativa* and soils allowed monitoring the genetic diversity of *Sinorhizobium meliloti* strains isolated from nodules formed over the years by the same single plant. The results showed that the genetic structure of the symbiotic rhizobial population was influenced, to different extents, by factors such as the soil type, plant cultivar, individual plants within the same cultivar and time. At the beginning of the experiment, the genetic variability of *S. meliloti* strains was mainly related to the soil and the host plant cultivar, while over the years, the effects of these two factors were no more significant, and the individual plant played the most important role [22]. Since the cultivars used also showed a relatively high genetic polymorphism [31], the individual plant effect retrieved could be attributed to either random genetic drift of the nodulating rhizobial populations or plant genotype selection.

Several pieces of evidence accumulated later confirmed that the symbiotic phenotype, including the amount of the bacteroid-differentiated population inside root nodules, is influenced by both the rhizobial strain and plant line [19]. Recently, the host nodule-specific PLAT domain (NPD) gene family, affecting late-stage nodule development and host biomass [30], as well as plant-nodule-specific Cys-rich peptides (NCRs), has been shown to have strain-specific effects on rhizobial fitness in competition [32]. The NPD gene family is composed of five genes that are expressed during late nodule development. NDP knockout mutants showed that the disruption of NDP genes led to changes in symbiotic phenotypes (using *S. meliloti* 1021 as a rhizobial symbiont), including plant size, number

of mature (pink) nodules, time of nodule senescence and the area of the fixation zone in nodules [30]. Interestingly, plants inoculated with a mixed population of rhizobial strains and with single strains showed phenotypic differences, suggesting that the effects of mutating NDP genes are rhizobial strain-specific [30]. Another important factor which can be related to partner choice is plant-encoded nodule cysteine-rich peptides (NCR peptides) involved in the formation of indeterminate nodules. NCR genes are exclusively expressed in infected nodule cells, demonstrating their relevant role as players in the regulation of symbiotic development. We discuss the relevance of NCR peptides later in this paper; however, here, it is important to point out their role in partner choice. The rhizobial symbiont in fact seems to be protected against the antibacterial activity of NCR peptides by a specific gene, *bacA* [15,33]. As a proof of this hypothesis, heterologous complementation of BacA from different rhizobial species produced host species-dependent symbiotic outcomes [34]. Moreover, an evolutionary scenario of host–symbiont coevolution with a rapid diversification of the NCR and *bacA* genes in legumes has been reported [33,34]. This evidence may suggest that allelic forms of the *bacA* gene may be involved in strain x host genotype specificities.

3. Fitness Alignment between Rhizobium and Host Plant

Symbiotic interactions between nitrogen-fixing rhizobia and host plants are a clear example of classical mutualism, where the increase in the population size of one partner causes an increase in its partner's population size. Rhizobium inoculation has a beneficial effect on legume growth, and the growth of a compatible host increases the relative abundance of rhizobia in the soil [35]. However, conflicts in symbiosis may arise in terms of conflicts of interest between the symbiont and the host and fitness conflict in the use of resources (cheating). A meta-analysis showed a widespread alignment between the host's and the symbiont's fitness [36]. Mutant studies have shown that a large number of mutations that are deleterious to the host are also deleterious to the symbiont's competitive ability, and, conversely, that beneficial mutations for the symbiont increase the host's growth. Cheater mutations, increasing rhizobium fitness at the host's expense, are rare. In fact, a trade-off between rhizobium fitness, use of resources and plant fitness exists. An integrated metabolic model reconstruction between the host plant *M. truncatula* and the symbiotic rhizobium *S. meliloti* has highlighted the fitness conflicts in the use of resources [37]. Here, increasing nodule biomass (therefore allowing more rhizobial cells to grow) initially increases host biomass, but after, a decrease in plant biomass gradually occurs. Considering nitrogen fixation efficiencies, simulations in the metabolic model indicated a linear relationship between nitrogen fixation and biomass production until an optimal level, but then excessive nitrogen fixation quickly resulted in a drop in plant biomass due to insufficient energy to support both the nitrogen fixation and ATP production costs. This pareto frontier between the nitrogen fixation efficiency and the rate of plant biomass production probably contributes to reducing cheating mutations, therefore aligning the symbiont and host fitness. However, we should consider that most of the studies measured plant fitness as plant growth/biomass production and did not directly measure reproductive capacity (e.g., seed production). Experiments directly assessing plant fitness would be needed to evaluate the alignment in fitness between the rhizobium and host plant in more depth. Moreover, directly measuring rhizobial cells released from nodules would also be relevant [38], in order to better understand the strategies of determinate and indeterminate nodule types, the latter harboring terminally differentiated bacteroids [36,39]. Moreover, most of the performed studies included single rhizobial strains, but competition between rhizobium strains in the formation of nodules exists, as well as evidence of variable levels of mutualisms in natural symbiont populations [40–43]. These give rise to the need for additional control over the mutualistic ability of the symbiont [44–46]. Experimental strategies, such as cycles of selection of a symbiont population, transposon insertion libraries, the use of massive genomic sequencing [47,48] and genome-wide association studies [41], allow shedding light on the genetic determinants for competition. These studies

have shown an incredible number of putative loci in the symbiont's genome, which can be related to competition and fitness in the indeterminate nodule colonization. Clarifying the importance of such loci in terms of their contribution to the host's fitness is compulsory, especially in relation to the use of rhizobia as inoculants in agriculture with the aim to increase legume crop yield [49]. Strains harboring large-scale genomic manipulations, such as those having a minimal symbiotic genome [50,51] and large genome manipulations [52], in addition to other more classical genetic approaches [20], would be useful to experimentally test predictions based on genome analyses and computational models.

4. Detecting Genotype × Genotype Interaction in Nodule Colonization

During the molecular dialog between legumes and rhizobia, several aspects come into play when the partner choice is to be made [53]. Phenotypic studies in legumes and rhizobia have shown that even within a species, symbionts are not all equivalent in the fitness benefits they confer to their hosts, and conversely, hosts differ in the fitness benefits they confer upon symbionts [54,55]. Sanctions are one of the possible strategies the host can choose to control the resource environment of their symbiont [56]. Legumes select for more cooperative rhizobia by imposing sanctions based on the amount of nitrogen fixed once bacteria are established inside nodules: effective symbionts are favored through the sanction of noneffective symbionts [57]. Several works have been conducted that attempt to understand the basis of sanctions and partner choice and the genetic determinants underlying preferences in host and symbiont partnerships [14,42,44,57–59]. However, only a few genes have been clearly identified as related to host preference. A classical element related to host preference is the *nod* genotype of the rhizobium, related to NF production. NFs are characterized by oligosaccharide chains of different lengths, distinct fatty acids and several substitutions, which determine the host range of the rhizobium, either among species or among strains of the same species (symbiovars) [60]. The *nod* genotype of rhizobia includes a set of genes encoding the enzymes that synthesize LCOs, whose basic structure is synthetized by proteins encoded by the *nodABC* genes [61]. The *nodC* gene is one of the most relevant variations in the LCO structure and a key determinant of species-specific recognition [62]. Interestingly, this recognition can be poorly specific, with a legume host being nodulated by different species/strains of rhizobia, showing extensive polymorphism in the *nodC* gene (e.g., *Glycine max*, *Lotus* spp., *Phaseolus vulgaris, Sophora flavescens*, *Arachis hypogaea*), or highly specific (e.g., *Cicer arietinum*, *Astragalus sinicus*, *Trifolium* spp.), requiring the presence of specific rhizobial species, with a highly conserved and distinct *nodC* gene sequence [62].

Alongside the now well-elucidated set of genes and functions related to the core symbiotic equipment of rhizobia (such as *nod* genes involved in Nod factor biosynthesis and flavonoid perception [63]), much remains unknown about the modifiers/accessory genes required to optimize the interaction. Genome-wide analyses of plant-associated microbes have found enrichment in genes related to carbohydrate metabolism, the type III secretion system, phytohormone production and phosphorus solubilization [8,64,65]. However, concerning the intraspecific selectivity of legume–rhizobia symbiosis, interaction with plant innate immunity seems to play an important role [66–68]. The interaction with plant immunity relies on various aspects of rhizobia (e.g., secretion systems, exopolysaccharides, proteolysis of plant antimicrobial peptides). Different rhizobial secretion systems (type III, type IV and type VI) influence rhizobial host specificity and the number of nodules, possibly modulating the transport of effector proteins into host cells [69]. Additionally, exopolysaccharides have been linked to infection and colonization processes by supporting bacterial attachment to the root surface. In rhizobia, secreted exopolysaccharides contribute to building a biofilm that increases nutrient absorption and enhances communication between plants and bacteria [70]. Early studies showed that mutation of genes related to the synthesis of exopolysaccharides (*exo* genes) in *S. meliloti* resulted in ineffective nodules on alfalfa, containing no bacteroids [71].

In general, the more studies are performed, especially coupling genome-wide mutant libraries with testing of symbiotic competitiveness, the more genes are found. These genes are involved in a plethora of functions, including metabolic pathways, transporters, chemotaxis and motility, from systems spanning from the *Bradyrhizobium diazoefficiens* host range to *S. meliloti* [48,72,73]. In addition to simple gene presence/absence on the rhizobial symbiont, recent genomic and transcriptomic studies have stirred the attention on genotype-by-genotype (G × G) interactions in symbiotic outcomes [41,49,52,74,75]. Experiments carried out on *M. truncatula* and *S. meliloti* strains demonstrated that the ability to enter plant nodules increased rhizobium fitness, and plants reward more cooperative rhizobium strains by preferentially forming nodules with them, but the rewarded strains differ among plant genotypes [45]. This evidence highlights how the outcomes of rhizobium fitness (hence their evolution) depend on the genetic composition of the plant populations they interact with.

Indeed, during the partner recognition nodule formation process, G × G interactions could act at different levels. In the early nodule developmental stage, i.e., during the reciprocal perception of the partners, symbiotic specificity is based on the rhizobial perception of root exudates presenting host-secreted flavonoids and the subsequent host perception of rhizobial Nod factors [76]. Plant exudates, key determinants of the rhizosphere microbiome structure, vary among plant species and different varieties of the same plant species [77]. By maintaining a broad repertoire of transporters, rhizosphere bacteria, including rhizobia, might efficiently capture resources from different types of hosts and successfully compete during plant colonization [64]. In *S. meliloti*, the pSymB chromid harbors many genes for membrane transporters, having roles in rhizosphere colonization and utilization of nutrients present in root exudates [78,79]. Deletion of pSymB has been shown to impair rhizosphere colonization [78,80]. Interestingly, the transcriptional patterns induced in *S. meliloti* by root exudates from different varieties of alfalfa changed according to the variety [55]. Moreover, different strains of *S. meliloti* responded differently to the same variety, clearly showing a G × G-related pattern of transcriptional response changes [55]. In the same work, a goodness-of-fit model (nested likelihood ratio test) indicated that nearly up to 30% of the total differentially expressed genes were explained by strain x variety interaction. Similar results were also obtained from the analysis of nodule transcriptomes (hence the mature phase of symbiosis) [54]. Among the retrieved genes, the main functions were cellular energy production pathways and processes playing roles in the exchange of nitrogen and carbohydrate utilization. Still concerning the nodule environment, during the shift from non-differentiated, reproductively capable rhizobia to the no longer reproductively competent and differentiated bacteroids, differential expression of host genes may represent mechanisms used by hosts to control differentiation and nitrogen fixation in rhizobia [81]. In this context, a group of genes encoding for nodule cysteine-rich peptides (NCR peptides) triggered interest for their possible role in modulating the outcome of symbiosis and as candidates for determining post-infection host–rhizobia specificity (Figure 1). Legume NCR peptides were initially discovered as antimicrobial molecules limiting the reproductive potential of endophytic rhizobia under sub-lethal concentrations [82,83]. However, they have been recently reconsidered as decisive for the bacteria to adapt to the intracellular environment, therefore being needed for a successful symbiosis [84,85]. On the other hand, some rhizobia are not challenged by NCR peptides during symbiosis, as NCR peptides are not present in their respective host legumes such as soybean and *Lotus* spp. [86]. This may be due to the disadvantage of legumes to impose terminal differentiation on bacteroids in legume species lacking persistent infection threads [86]. However, a comprehensive interpretation of the reason NCR peptides are essential for symbiotic nitrogen fixation in some legume species and not in others is still lacking [86]. Screening of NCR peptide genes from several *M. truncatula* accessions reported significant expression and sequence variation, suggesting they can be involved in symbiotic phenotypic variation among accessions [87]. Indeed, NCR peptides are responsible for many of the metabolic and morphological changes observed in bacteroids in nodules [33]. In a classical Red

Queen evolutionary scenario [88], a rhizobial-encoded peptidase, termed as host range restriction peptidase (HrrP), was shown to allow the modulation of the *S. meliloti* host range toward different *Medicago* species by specific proteolysis of host-derived NCR peptides [89] (Figure 1). Experiments of transposon mutagenesis followed by high-throughput sequencing (Tn-seq) have identified genes, in addition to HrrP, that increase or decrease bacterial competitiveness during exposure to the NCR peptides in *S. meliloti*. These genes relate to several cellular processes such as polysaccharide biosynthesis, membrane proteins, peptidoglycan effector proteins, transcriptional regulators and ribosome-associated factors. It is interesting to note that for a single component of nodule colonization, such as the NCR sensitivity, as well as for the early interaction with roots, a panoply of genes is involved.

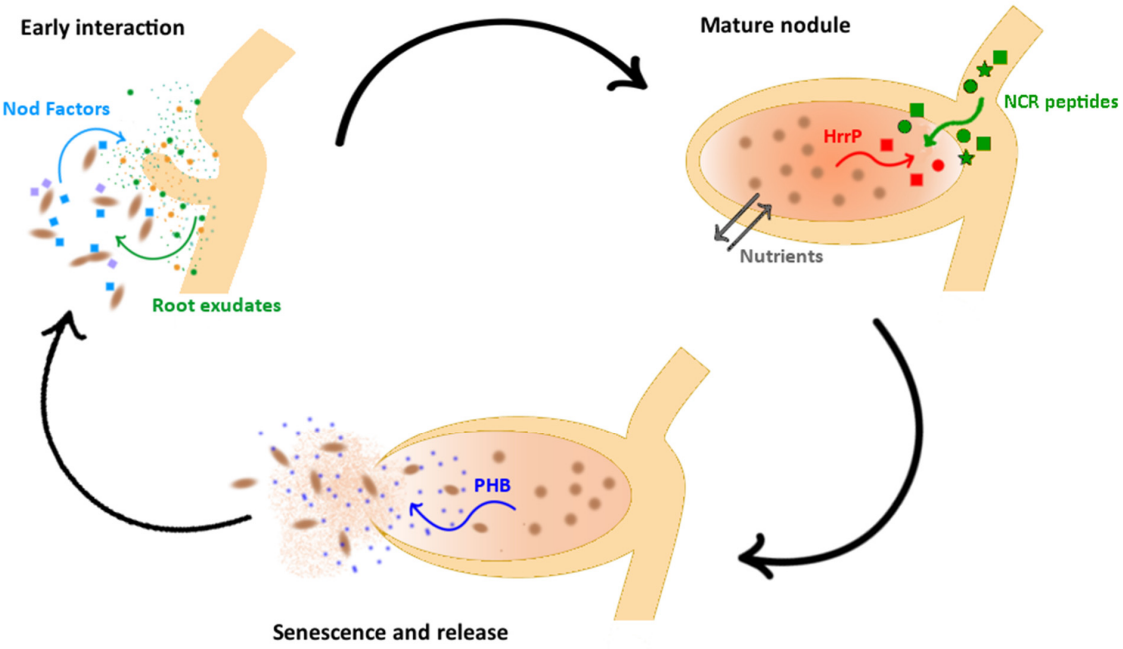


Figure 1. Genotype × genotype (G × G) interactions in the nitrogen-fixing root nodule symbiosis. A schematic example of three phases of symbiosis, namely, early interaction with roots, mature nodules and release of bacteria from dehiscent nodules, is shown. The key molecular actors identified along the phases are reported.

Among the sanctions the plant can induce on nodules, senescence related to ethylene production has also been discussed. Strains of rhizobia expressing the 1-aminocyclopropane-1-carboxylate (ACC) deaminase enzyme, a possible scavenger of the ethylene precursor ACC, have been shown to have, in some cases, advantages for nodule colonization [40], though in *S. meliloti*, the gene *acdS*, encoding for ACC deaminase, also seems to be related to rhizosphere colonization and the use of unusual nitrogen sources [90]. Manipulation of ethylene by rhizobia is also due to rhizobitoxine [91]. Strains producing rhizobitoxine are more competitive over nonproducing strains and also accumulate more carbon resources, such as poly-β-hydroxybutyrate (PHB) [92]. PHB accumulation inside bacteroids is claimed to be a measure of rhizobial fitness, once the dehiscent (determinate) nodules release the bacteria [38] (Figure 1). PHB stabilizes the cellular redox condition and provides a carbon source for bacteroids/undifferentiated rhizobia when released from the nodule to the soil [93,94]. While some strains (e.g., *Sinorhizobium* sp. NGR234, *B. elkani* USDA 76, *Sinorhizobium meliloti* CCBAU01199) produce a large amount of PHB, others (e.g., *M. loti* NZP2213, *B. arachidis* CCBAU 051107, *M. septentrionale* CCBAU03399) produce only a few PHB granules [95]. It has been shown that a single rhizobial population may display bet-edging on PHB, with two subpopulations, one with low PHB which has greater competitiveness for resources, and another with high PHB which can survive until a legume host

is next available [93,96]. However, the role of genetic variability in ethylene modulation and PHB production in connection with plant metabolism (a G × G interaction) has not been addressed yet.

We can argue that many molecular actors in modulating G × G interaction in symbiosis are still unknown, and we need a more comprehensive approach, using systems biology simulation and multi-omics data interpretation, combined with careful phenotypic analyses, to fully disclose the network of relationships among rhizobial and plant genes.

5. Molecular Tasting in the Rhizosphere

Bacterial root colonization (on the rhizoplane and in the endosphere) is stimulated by the panoply of molecules released by the plant root tissues, either as volatile or soluble molecules. Concerning volatile molecules (often referred to as volatile organic compounds, VOCs), relatively few are known in relation to direct bacterial chemoattraction. Only recently has a role of volatile compounds in recruiting beneficial environmental bacteria, for instance, under stress conditions (the "crying-for-help" strategy [97]) been demonstrated [98–100]. However, no specific role of plant-emitted VOCs has been assigned to the symbiosis with rhizobia. On the contrary, for rhizobium-emitted VOCs, there is evidence that they may have a role in the promotion of non-host plant growth and in the increase in iron uptake mechanisms, rhizosphere acidification and increased root ferric reductase in host legumes [101]. However, most of the research and data are on soluble molecules, emitted by roots and microbes. Concerning the microbial contribution, very detailed models of bacterial population density control and adhesion to plant roots via biofilm formation are available [102]. Quorum sensing as microbe–microbe signaling is a well-known phenomenon mediated by molecules such as N-acyl homoserine lactones (AHLs) [102]. In rhizobia, quorum sensing systems are related to exopolysaccharide production and early root colonization [103], and to symbiotic plasmid transfer in some species such as *R. leguminosarum* bv. *viciae* (pRL1JI) and *Sinorhizobium fredii* NGR234 (pNGR234) [104]. Interestingly, several compounds from the plant side (present in root exudates) are agonists of AHL signaling [105], indicating that host plants can manipulate the bacterial behavior, possibly favoring the expression of bacterial genes that are beneficial to them. Indeed, root exudates are a very complex mixture of substances, which vary depending on the plant genotype, plant growth stage and substrate, dramatically affecting root microbiota [106–108]. Legume root exudates include sugars, organic acids and amino acids in large amounts, but also molecules such as flavones and flavonoids, well-known inducers of nodulation signaling [109–111]. These compounds elicit various functions in the rhizobia present in the rhizosphere, spanning from metabolic modules [80] to transcriptional changes in genes related to chemotaxis and transport [112], and also mirroring competitive abilities [113]. In a model of G × G interaction between alfalfa (*Medicago sativa*) varieties and strains of the rhizobium *S. meliloti* [74], different plant varieties affected the transcription of different sets of rhizobial genes, including those involved in quorum sensing systems and motility. Since the composition of root exudates among alfalfa varieties mainly differs in the content of amino acids, such as N-acetyl-L-leucine, tryptophan, cytosine, 3,5-dihydroxyphenylglycine and some dipeptides [74], we can hypothesize that some of these molecules may play a signaling role with respect to symbiotic rhizobia. In line with hypotheses on the presence of molecules which can differentially modulate rhizobia gene expression and behavior in root exudates, there is evidence that root exudates from non-host plant species (such as wheat) may synergistically act to promote rhizobia–host legume symbiotic interaction [114].

Of course, the most well-known signaling molecules present in legume root exudates are flavonoids such as luteolin, quercetin and apigenin. Flavonoids modulate the host specificity of symbiotic rhizobia [115], selectively binding and activating the protein NodD, which transcriptionally regulates the so-called nodulation (*nod*) genes. These are a set of three common (*nodA*, *nodB*, *nodC*) genes and various other genes which encode the enzymes responsible for the synthesis and decoration of lipochitooligosaccharide molecules known as Nod factors. Nod factors differ among rhizobial species/strains and, once released by the

bacteria, bind receptor proteins on the host plant, allowing a species-specific recognition, prior to the formation of the symbiotic nodule [12]. Diverse flavonoids may act as either an agonist or an antagonist of the same NodD protein [116], and different NodD proteins may differ in the specific agonist and antagonist. This extreme selectivity may explain part of the variable range of the transcriptional response to the same flavonoid (e.g., luteolin) by different strains of rhizobia [74]. However, flavonoids affect functions other than solely *nod* gene transcription. For instance, luteolin was shown to modulate the expression of motility genes in the *S. meliloti* 1021 strain [74]. In the same work, the presence of G × G interactions between three varieties of the host legume *M. sativa* and three rhizobia strains of *S. meliloti* was tested as the transcriptomic response of bacterial strains to root exudates. Interestingly, root exudates from the three plant varieties influenced a core set of functional gene categories in all three bacterial strains, and most of them were related to protein expression and modification, carbohydrate transport and metabolism and energy production and conservation [74]. Recent studies showed that luteolin can also modulate various substrate utilization pathways, as well as eliciting resistance toward stressors, including several antibiotics, toxic ions, respiration inhibitors, membrane damagers, DNA intercalants and other antimicrobial agents [117]. In other studies, it has been demonstrated that flavonoids can affect rhizobial chemotaxis by acting as the main attractant molecules prior to the initial signaling stage [118], and they can influence the production of proteins released in the soil from rhizobia via the type I and III secretion systems and the production of polysaccharides localized on the bacterial surface [119].

6. Role of the Rhizobial Mobile Genome in the Perception of the Plant

Several rhizobial species harbor a multipartite genome, i.e., a genome with a divided structure in several replicons, namely, a chromosome and various chromids, megaplasmids and plasmids [120]. Previous works and reviews have discussed the advantages of having a multipartite genome structure, such as maintaining a fast duplication rate while increasing the genome size and functional organization [78,120]. This genome organization is relatively frequent in bacterial species which interact with hosts and may colonize different environments, such as rhizobia. In rhizobia with a multipartite genome structure (such as those belonging to the genera *Sinorhizobium* and *Rhizobium*), genes encoding functions related to symbiosis and nitrogen fixation are mostly located on megaplasmids [121–123]. Such megaplasmids have clear signatures (such as the GC% content, codon usage and insertion sequences, IS) of alien origin with respect to the chromosome and likely spread relatively recently through horizontal gene transfer in rhizobia, giving rise to the evolution of the symbiotic phenotype [79,124]. Symbiotic megaplasmids represent classical examples of large mobile genetic elements (MGEs), providing novel and complex traits (i.e., requiring many genes), such as interactions with eukaryotic hosts [125]. In the rhizobial strain *S. meliloti* 1021, the symbiotic megaplasmid (pSymA ~1.4 Mb in size) represents nearly 21% of the genome and carries the genes critical for the symbiosis process such as *nod*, *nif* and *fix* [126]. Genetic experiments with mutants lacking the pSymA megaplasmid have clearly shown that this element is dispensable, but essential for symbiosis. Strains lacking pSymA are still able to thrive in the soil and colonize the rhizosphere but cannot allow the formation of symbiotic nodules [78,127]. Since pSymA contains genes required for plant recognition (e.g., *nod* genes), shuffling of the pSymA megaplasmid led to changes in the symbiotic phenotype [52]. Concerning plant signal recognition in particular, *nod* gene diversity can determine large variation in the host range. The term "symbiovar", mentioned above, was introduced in the nomenclature of rhizobia aiming to reflect such an adaptation to different host species by the strains of the same rhizobial species [60]. This term emphasizes the spreading by horizontal gene transfer (HGT) of symbiotic plasmids and symbiotic-related genes. In this context, it should be noted that symbiotic megaplasmids often contain a number of mobile genetic elements, such as IS and mobile introns, which determine large structural variation [79]. Symbiotic megaplasmids harbor genes other than *nod*, *nif* and *fix* that are relevant in the interaction with the host plant and can provide different outcomes

of symbiotic phenotypes. Among the most discussed is the gene encoding the *acdS* gene, which is claimed to be relevant for scavenging the ethylene precursor ACC, allowing a local decrease in ethylene and consequently nodule senescence [128]. However, the role of *acdS* is controversial, since the same gene of *S. meliloti* also appears to be expressed in the rhizosphere of non-host plants [90]. Other functions, possibly related to nitrogen metabolism, but also potentially involved in resistance to harsh environments, are located on pSymA, such as the gene encoding the nickel/proton antiporter *nreB* [129]. Interestingly, both *acdS* and *nreB* display evidence of HGT, uncoupled from the pSymA phylogeny, emphasizing the relevance of symbiotic megaplasmids as hot spots for the acquisition and spread of the rhizobial mobilome.

Though the symbiotic megaplasmids, such as pSymA of *S. meliloti* 1021, confer the ability to become a symbiont, they are not simple "plug-and-play" modules, requiring a regulatory and metabolic connection to the other resident replicons of the genome (in *S. meliloti* 1021, the chromosome and the pSymB chromid) (Figure 2). Simulations of regulatory networks predicted the existence of several cross-replicon interactions [130], indicating that symbiotic megaplasmids need to be wired to the rhizobial genome. Transposon sequencing (Tn-Seq) experiments carried out on *S. meliloti* strains lacking pSymA and pSymB indicated that almost 10% of the chromosomal genes interact with genes located on the symbiotic megaplasmid pSymA and the pSymB chromid [131]. Concerning pSymB, several functions related to the exploitation of the rhizosphere, such as transport of carbohydrates, are present, suggesting a role in the differential exploitation of plant genotype-specific carbon sources in the rhizosphere [80,121]. In line with this hypothesis, transcriptome analysis of *S. meliloti* strains incubated with root exudates from different alfalfa varieties showed that pSymB stimulons (the set of differentially expressed genes after stimulation with root exudates) are highly variable either among strains or among conditions [74]. Among strains, stimulons ranged from 17% to 24% of the total differentially expressed genes, and among conditions (root exudates from different plant varieties and luteolin as a model flavonoid for *nod* gene induction), they ranged from 13% to 35% in the same strain [74], indicating that pSymB-harbored genes play a relevant role in the genotypic interaction with host plants.

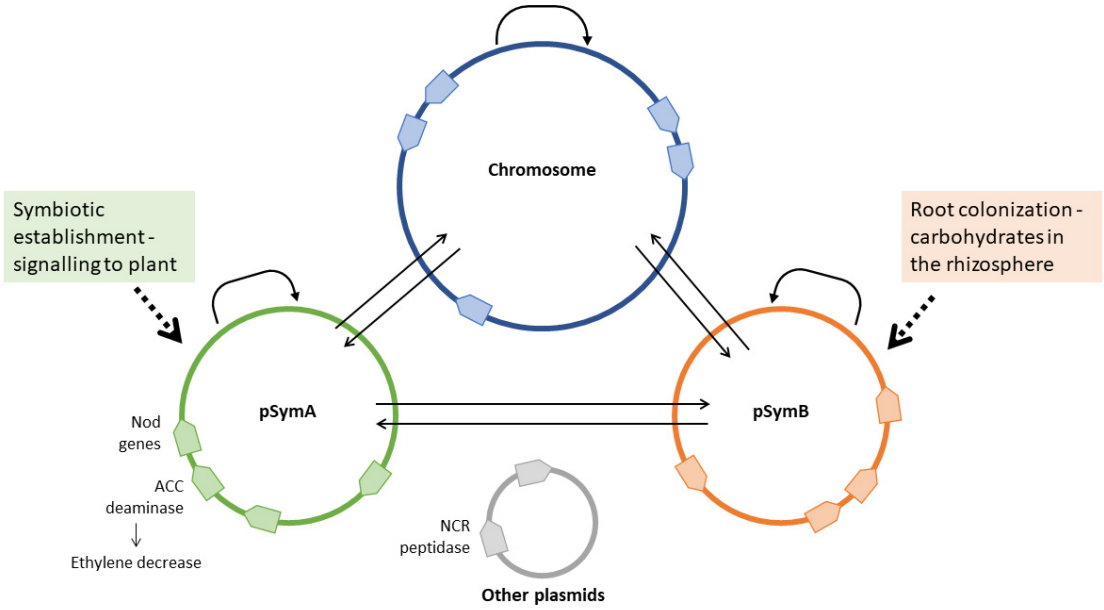


Figure 2. Partition and interaction among replicons in host genotype-specific interaction. The example reported refers to the *Sinorhizobium meliloti* genome, composed of a chromosome, a chromid (pSymB), a megaplasmid (pSymA) and often smaller accessory plasmids. Arrows connecting replicons are based on gene interaction data from [74,130,131].

Clearly, megaplasmids containing symbiosis genes, but also other plasmids (such as small accessory plasmids which may harbor NCR peptidases) and chromids, can be transferred to other rhizobia, promoting population evolution and the emergence of strains with a novel host genotype preference. Plasmid transfer is often controlled by quorum sensing, with high bacterial cell densities inducing transfer [132,133]. High cell densities could hardly be present in bulk soil, while they can likely be achieved in the rhizosphere, suggesting a role of the host in influencing the spread of symbiosis genes, similar to *Agrobacterium tumefaciens* pTi plasmid transfer [134]. However, though genetic dissection of conjugation genes has been performed, few data are available on the environmental signals and context of the symbiotic plasmid conjugal transfer [135–138]. Besides plasmids, within the rhizobia genome, symbiotic regions such as integrative and conjugative elements (ICEs [139]), determining host interaction and subjected to HGT, are present. Interestingly, an in vitro experiment showed that certain plant compounds inducing the nodulation process also improve the transfer of symbiosis islands in *Azorhizobium caulinodans* [140]. Here, the symbiosis island of 87.6 kb is an ICE capable of transferring to a site-specific gly-tRNA gene of other rhizobial species, expanding their host range. An AhaR transcriptional regulatory protein of the LysR family, located in the ICE, triggers the HGT process in response to plant flavonoids, which also induces the expression of nodulation genes through NodD. We can speculate that symbiotic megaplasmid transfer is controlled by similar environmental signals related to the rhizosphere or endosphere of host plants. Experiments clarifying these signals are needed to better understand the ecological context of symbiotic determinants spreading in the rhizobial populations.

7. A Cross-Talk in the Rhizomicrobiota

The more we delve deep into the complexity of symbiotic interactions, the more we discover that the overall biota surrounding plant roots has a relevance in the establishment of the success of symbiosis and of the phenotypes (i.e., performances) of both partners. It is becoming clear that many more signaling molecules than those initially discovered are present, including VOCs. These molecules play a role in both intermicrobial and interkingdom interactions. The acknowledgement of their role in microbe–microbe interactions is quite recent. Some of them exert antimicrobial activity, while others alter cell motility or induce biofilm formation, but little is known about the molecular mechanisms of these responses. On the other hand, their role in the interaction with plants has been known for almost two decades, and their emission improves plant growth by stimulating root development, chlorophyll production and the uptake of essential elements such as iron. They are also involved in plant tolerance to abiotic and biotic stresses, through antibiotic activity and activation of plant immunity [101].

Works on *Arabidopsis thaliana* have nicely disclosed the many actors and levels of interaction in the plant–microbe system [141,142]. Roots of a healthy plant host a variety of microorganisms belonging to different kingdoms, such as bacteria, archaea, protists and fungi (Figure 3). Bacterial communities are taxonomically structured, and their assembly is driven by factors such as soil type, plant compartment, host genotype/species and plant immune system. Fungal communities, in contrast, seem to be more influenced by their biogeographical distribution. Regarding protists, little information is available about community profiles. Among members of the plant-associated microbiota, two types of interactions can be identified: cooperation and competition. Microorganisms employ several cooperative mechanisms, such as nutritional interdependencies, biofilm formation, molecular communication through quorum sensing, enhanced dispersal and bacterial endosymbiosis in fungi, in order to ensure their persistence within the plant holobiont. Competitive mechanisms include resource competition through rapid utilization of limiting resources and secretion of siderophores, contact-dependent competition through the bacterial type VI and type III secretion systems, secretion of antibacterial and antifungal compounds and emission of VOCs produced by bacteria and predation [141]. For the predation aspect, we do not know if and how host legume plants perceive VOC signals

from rhizobia and if, in turn, a plant which is forming nitrogen-fixing nodules emits VOCs to interact with microbes, other plants or pollinating insects. Setting up in vitro and in vivo experiments where VOCs are measured and identified using highly sensitive techniques such as proton transfer reaction-time of flight-mass spectrometry (PTR-TOF-MS) could allow shedding light on this still unknown side of rhizobia–host interaction.

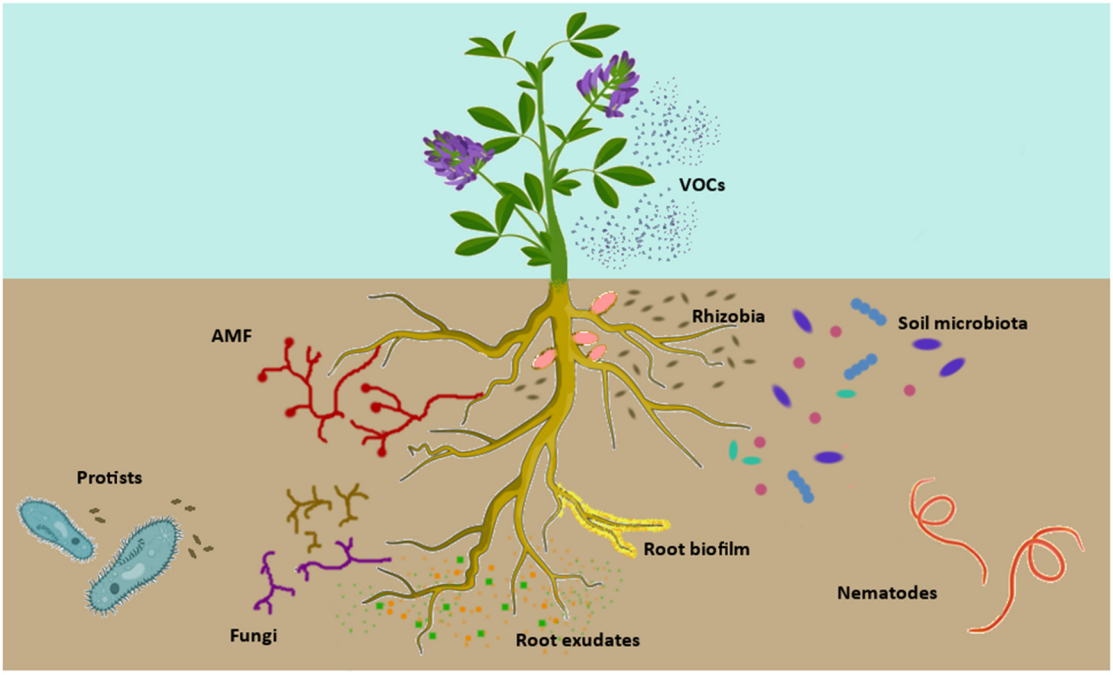


Figure 3. Cross-talk in the legume rhizosphere. Within- and among-kingdom interactions giving rise to modulation of symbiotic establishment are depicted, spanning from protists' role as grazers of rhizobia to fungi, soil microbiota and nematodes which are attracted toward roots and may transport rhizobia.

For legumes, several clues on the interaction between rhizobia and fungi, especially arbuscular mycorrhizal fungi, are present (for a recent review, see [143]). Arbuscular mycorrhizal fungi (AMF) penetrate the cortical cells of the roots and ensure nutritional uptake to the plant. Mycorrhiza formation is guided by a complex signal exchange between the host plant and the AM fungus [144]. Strigolactone secreted by the plant induces spore germination, hyphal branching and hyphal growth toward the roots. This leads to the synthesis of short-chain chitin oligomers (CO4 and CO5), structurally similar to the rhizobial Nod factors. Recently, the receptors discriminating AMF and rhizobial signaling have been discovered and characterized [145]. Indeed, it is clear that AMF–plant symbiosis predates rhizobial–legume symbiosis, the latter being evolved from a previously settled molecular dialog. The first observation of AMF effects on nodulation and plant growth dates back almost fifty years [145]. Later on, many other beneficial effects, such as enhancement of symbiotic nitrogen fixation in drought soils, and the increase in the number and dry weight of nodules, have been observed [145]. In the establishment of the tripartite symbiosis among legumes, AMF and rhizobia, one study [146] highlighted the importance of specific flavonoids such as daidzein, genistein and coumestrol which act as signal molecules, playing a key role in the early stages of the interaction. This tripartite interaction represents a multiple mutualism where a host interacts with two mutualistic species. A recent study [147] showed that mycorrhizal fungi have a strong impact on fitness alignment between *M. truncatula* and rhizobia; in fact, in the presence of mycorrhizal fungi, rhizobia and plant fitness was more positively aligned and the strength of selection of a host trait (AMF branch number) doubled compared to plants inoculated with rhizobia alone.

Further dissection of the rhizospheric cross-talk is therefore crucial if we want to better understand the role of microbial networks in holobiont fitness, in view of agronomic applications [143]. Here, as mentioned before, a systems biology approach and more realistic experimental setups can help to overcome the symbiont- or host-centric single visions and clarify the effects of such symbioses on all partners [141,147].

8. Future Directions

A large array of molecular signals is exchanged between rhizobia and the host legume plant; some of them are crucial for the symbiosis to take place (such as flavonoids and Nod factors), while others are modifiers of the interaction. The latter have great importance in competition in the microbiota, in genotype-specific perception of host plants by the rhizobium and in the ability of the plant to maximize symbiont performances (i.e., nitrogen fixation and metabolic exchange) [143]. Moreover, the fungal component of the microbiota is stirring increased attention. Representing more than 70% of all land plants, legumes are able to directly interact with arbuscular mycorrhizal fungi (AMF), and direct molecular dialog is present between AMF and rhizobia [148]. However, we still do not know the molecular interactions rhizobia have with other key components of the soil biota, such as fungi other than AMF which are relevant components of bioinoculants used in agriculture (e.g., *Trichoderma*), or protists and nematodes, which can have an effect on the rhizobia population size (e.g., through grazing on rhizobia), but also the mobility of rhizobia toward the roots (e.g., nematode-carrying rhizobia being attracted by plant-released volatiles) [149].

A relatively simple problem, namely, rhizobium–plant symbiosis, is becoming very intricate, with many more genetic modifiers of the symbiosis than initially thought residing in the genomic variability of the partners, but also in a partnership which is not between two partners—the plant and the rhizobium—but with the overall soil biota. The possibility to exploit strains and genotypes for designing tailored bioinoculants, able to increase crop yield, must consider such intricate interactions. Consequently, we need to develop model systems, such as novel in vitro experimental setups and data interpretative models, to predict the outcomes from lab-scale analyses and trials to open field applications [150]. Systems biology approaches have proved to be powerful in finding key parameters in complex biological phenomena and have already shown their ability in modeling rhizobia symbiotic interaction [37,151,152]. Ecological modeling of plant–rhizobium–soil biota interactions [153], coupled with molecular data coming from genome analyses (see, for instance, [154] concerning human–microbe interaction), should be prioritized in order to find predictors (especially in the rhizobia genomes) of the goodness of the symbiotic phenotype in nature (and field conditions).

Author Contributions: Conceptualization, C.F., A.M. and L.C.; writing—original draft preparation, C.F., F.V., M.F., A.M. and C.F.; writing—review and editing, C.F., F.V., M.F., A.M. and C.F.; visualization, C.F.; supervision, A.M.; funding acquisition, A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the grant MICRO4Legumes, D.M. n. 89267 (Italian Ministry of Agriculture) to A.M. L.C. is supported by a PhD fellowship from MICRO4Legumes. F.V. is supported by a PhD fellowship from PON Research and Innovation 2014–2020, D.M. n. 1061/2021 (Italian Ministry of University and Research). C.F. is supported by a post-doctoral fellowship from the H2020 ERA-NETs SUSFOOD2 and CORE Organic Cofund, under the Joint SUSFOOD2/CORE Organic Call 2019.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. O'Brien, A.M.; Jack, C.N.; Friesen, M.L.; Frederickson, M.E. Whose trait is it anyways? Coevolution of joint phenotypes and genetic architecture in mutualisms. *Proc. R. Soc. B Biol. Sci.* **2021**, *288*, 20202483. [CrossRef] [PubMed]
2. Wagner, M.R. Prioritizing host phenotype to understand microbiome heritability in plants. *New Phytol.* **2021**, *232*, 502–509. [CrossRef] [PubMed]

3. Friesen, M.L.; Porter, S.S.; Stark, S.C.; Von Wettberg, E.J.; Sachs, J.L.; Martinez-Romero, E. Microbially mediated plant functional traits. *Annu. Rev. Ecol. Evol. Syst.* **2011**, *42*, 23–46. [CrossRef]
4. Batstone, R.T. Genomes within genomes: Nested symbiosis and its implications for plant evolution. *New Phytol.* **2022**, *234*, 28–34. [CrossRef] [PubMed]
5. Bacci, G.; Cerri, M.; Lastrucci, L.; Ferranti, F.; Ferri, V.; Foggi, B.; Gigante, D.; Venanzoni, R.; Viciani, D.; Mengoni, A.; et al. Applying predictive models to decipher rhizobacterial modifications in common reed die-back affected populations. *Sci. Total Environ.* **2018**, *642*, 708–722. [CrossRef] [PubMed]
6. Hou, S.; Thiergart, T.; Vannier, N.; Mesny, F.; Ziegler, J.; Pickel, B.; Hacquard, S. A microbiota–root–shoot circuit favours Arabidopsis growth over defence under suboptimal light. *Nat. Plants* **2021**, *7*, 1078–1092. [CrossRef] [PubMed]
7. do Amaral, F.P.; Tuleski, T.R.; Pankievicz, V.C.S.; Melnyk, R.A.; Arkin, A.P.; Griffitts, J.; Tadra-Sfeir, M.Z.; Maltempi de Souza, E.; Deutschbauer, A.; Monteiro, R.A.; et al. Diverse Bacterial Genes Modulate Plant Root Association by Beneficial Bacteria. *mBio* **2020**, *11*, e03078-20. [CrossRef] [PubMed]
8. Levy, A.; Gonzalez, I.S.; Mittelviefhaus, M.; Clingenpeel, S.; Paredes, S.H.; Miao, J.; Wang, K.; Devescovi, G.; Stillman, K.; Monteiro, F.; et al. Genomic features of bacterial adaptation to plants. *Nat. Genet.* **2018**, *50*, 138–150. [CrossRef]
9. Levy, A.; Conway, J.M.; Dangl, J.L.; Woyke, T. Elucidating Bacterial Gene Functions in the Plant Microbiome. *Cell Host Microbe* **2018**, *24*, 475–485. [CrossRef] [PubMed]
10. Bulgarelli, D.; Garrido-Oter, R.; Münch, P.C.; Weiman, A.; Dröge, J.; Pan, Y.; McHardy, A.C.; Schulze-Lefert, P. Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* **2015**, *17*, 392–403. [CrossRef]
11. Poole, P.; Ramachandran, V.; Terpolilli, J. Rhizobia: From saprophytes to endosymbionts. *Nat. Rev. Microbiol.* **2018**, *16*, 291–303. [CrossRef] [PubMed]
12. Oldroyd, G.E.D.; Murray, J.D.; Poole, P.S.; Downie, J.A. The rules of engagement in the legume-rhizobial symbiosis. *Annu. Rev. Genet.* **2011**, *45*, 119–144. [CrossRef] [PubMed]
13. Hawkins, J.P.; Oresnik, I.J. The Rhizobium-Legume Symbiosis: Co-opting Successful Stress Management. *Front. Plant Sci.* **2022**, *12*, 796045. [CrossRef]
14. Ledermann, R.; Schulte, C.C.M.; Poole, P.S. How Rhizobia Adapt to the Nodule Environment. *J. Bacteriol.* **2021**, *203*, e00539-20. [CrossRef] [PubMed]
15. Nicoud, Q.; Barrière, Q.; Busset, N.; Bourge, M.; Le Bars, R.; Mergaert, P. *Sinorhizobium meliloti* functions required for resistance to antimicrobial NCR peptides and bacteroid differentiation. *mBio* **2021**, *12*, e00895-21. [CrossRef] [PubMed]
16. Haag, A.F.; Arnold, M.F.F.; Myka, K.K.; Kerscher, B.; Dall'Angelo, S.; Zanda, M.; Mergaert, P.; Ferguson, G.P. Molecular insights into bacteroid development during Rhizobium–legume symbiosis. *FEMS Microbiol. Rev.* **2013**, *37*, 364–383. [CrossRef] [PubMed]
17. Kereszt, A.; Mergaert, P.; Kondorosi, E. Bacteroid development in legume nodules: Evolution of mutual benefit or of sacrificial victims? *Mol. Plant-Microbe Interact.* **2011**, *24*, 1300–1309. [CrossRef]
18. Crook, M.B.; Lindsay, D.P.; Biggs, M.B.; Bentley, J.S.; Price, J.C.; Clement, S.C.; Clement, M.J.; Long, S.R.; Griffitts, J.S. Rhizobial Plasmids That Cause Impaired Symbiotic Nitrogen Fixation and Enhanced Host Invasion. *Mol. Plant-Microbe Interact.* **2012**, *25*, 1026–1033. [CrossRef] [PubMed]
19. Kazmierczak, T.; Nagymihaly, M.; Lamouche, F.; Barriére, Q.; Guefrachi, I.; Alunni, B.; Ouadghiri, M.; Ibijbijen, J.; Kondorosi, É.; Mergaert, P.; et al. Specific host-responsive associations between *Medicago truncatula* accessions and *Sinorhizobium* strains. *Mol. Plant-Microbe Interact.* **2017**, *30*, 399–409. [CrossRef]
20. Dicenzo, G.C.; Zamani, M.; Checcucci, A.; Fondi, M.; Griffitts, J.S.; Finan, T.M.; Mengoni, A. Multidisciplinary approaches for studying rhizobium–legume symbioses. *Can. J. Microbiol.* **2019**, *65*, 1–33. [CrossRef] [PubMed]
21. Roy, S.; Liu, W.; Nandety, R.S.; Crook, A.D.; Mysore, K.S.; Pislariu, C.I.; Frugoli, J.A.; Dickstein, R.; Udvardi, M.K. Celebrating 20 years of genetic discoveries in legume nodulation and symbiotic nitrogen fixation. *Plant Cell* **2020**, *32*, 15–41. [CrossRef] [PubMed]
22. Carelli, M.; Gnocchi, S.; Fancelli, S.; Mengoni, A.; Paffetti, D.; Scotti, C.; Bazzicalupo, M. Genetic Diversity and Dynamics of *Sinorhizobium meliloti* Populations Nodulating Different Alfalfa Cultivars in Italian Soils. *Appl. Environ. Microbiol.* **2000**, *66*, 4785–4789. [CrossRef]
23. Paffetti, D.; Scotti, C.; Gnocchi, S.; Fancelli, S. Genetic Diversity of an Italian *Rhizobium meliloti* Population from Different Medicago sativa Varieties. *Appl. Environ. Microbiol.* **1996**, *62*, 2279–2285. [CrossRef]
24. Kumar, N.; Lad, G.; Giuntini, E.; Kaye, M.E.; Udomwong, P.; Shamsani, N.J.; Young, J.P.W.; Bailly, X. Bacterial genospecies that are not ecologically coherent: Population genomics of *Rhizobium leguminosarum*. *Open Biol.* **2015**, *5*, 140133. [CrossRef] [PubMed]
25. Bailly, X.; Giuntini, E.; Sexton, M.C.; Lower, R.P.J.; Harrison, P.W.; Kumar, N.; Young, J.P.W. Population genomics of *Sinorhizobium medicae* based on low-coverage sequencing of sympatric isolates. *ISME J.* **2011**, *5*, 1722–1734. [CrossRef] [PubMed]
26. Cui, W.-J.; Zhang, B.; Zhao, R.; Liu, L.-X.; Jiao, J.; Zhang, Z.; Tian, C.-F. Lineage-Specific Rewiring of Core Pathways Predating Innovation of Legume Nodules Shapes Symbiotic Efficiency. *mSystems* **2021**, *6*, e01299-20. [CrossRef] [PubMed]
27. Epstein, B.; Branca, A.; Mudge, J.; Bharti, A.K.; Briskine, R.; Farmer, A.D.; Sugawara, M.; Young, N.D.; Sadowsky, M.J.; Tiffin, P. Population Genomics of the Facultatively Mutualistic Bacteria *Sinorhizobium meliloti* and S. *medicae*. *Genomics* **2012**, *8*, e1002868. [CrossRef] [PubMed]
28. Jorrin, B.; Imperial, J. Population genomics analysis of legume host preference for specific rhizobial genotypes in the *Rhizobium leguminosarum* bv. *viciae* symbioses. *Mol. Plant-Microbe Interact.* **2015**, *28*, 310–318. [CrossRef] [PubMed]

29. Bailly, X.; Olivier, I.; De Mita, S.; Cleyet-Marel, J.-C.; Béna, G. Recombination and selection shape the molecular diversity pattern of nitrogen-fixing *Sinorhizobium* sp. associated to *Medicago*. *Mol. Ecol.* **2006**, *15*, 2719–2734. [CrossRef] [PubMed]
30. Burghardt, L.T.; Trujillo, D.I.; Epstein, B.; Tiffin, P.; Young, N.D. A Select and Resequence Approach Reveals Strain-Specific Effects of *Medicago* Nodule-Specific PLAT-Domain Genes. *Plant Physiol.* **2020**, *182*, 463–471. [CrossRef] [PubMed]
31. Mengoni, A.; Gori, A.; Bazzicalupo, M. Use of RAPD and microsatellite (SSR) variation to assess genetic relationships among populations of tetraploid alfalfa, Medicago sativa. *Plant Breed.* **2000**, *119*, 311–317. [CrossRef]
32. Wang, Q.; Liu, J.; Li, H.; Yang, S.; Kereszt, A.; Zhu, H. Nodule-Specific Cysteine-Rich Peptides Negatively Regulate Nitrogen-Fixing Symbiosis in a Strain-Specific Manner in Medicago truncatula. *Mol. Plant-Microbe Interact.* **2018**, *31*, 240–248. [CrossRef] [PubMed]
33. Haag, A.F.; Baloban, M.; Sani, M.; Kerscher, B.; Pierre, O.; Angelo, S.D.; Kondorosi, E.; Longhi, R.; Boncompagni, E.; He, D.; et al. Protection of *Sinorhizobium* against Host Cysteine-Rich Antimicrobial Peptides Is Critical for Symbiosis. *PLoS Biol.* **2011**, *9*, e1001169. [CrossRef] [PubMed]
34. DiCenzo, G.C.; Zamani, M.; Ludwig, H.N.; Finan, T.M. Heterologous Complementation Reveals a Specialized Activity for BacA in the *Medicago—Sinorhizobium meliloti* Symbiosis. *Mol. Plant Microbe Interact.* **2017**, *30*, 312–324. [CrossRef] [PubMed]
35. Denison, R.F.; Kiers, E.T. Life histories of symbiotic rhizobia and mycorrhizal fungi. *Curr. Biol.* **2011**, *21*, R775–R785. [CrossRef] [PubMed]
36. Friesen, M.L. Widespread fitness alignment in the legume-rhizobium symbiosis. *New Phytol.* **2012**, *194*, 1096–1111. [CrossRef] [PubMed]
37. diCenzo, G.C.; Tesi, M.; Pfau, T.; Mengoni, A.; Fondi, M. Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. *Nat. Commun.* **2020**, *11*, 2574. [CrossRef] [PubMed]
38. Ratcliff, W.C.; Underbakke, K.; Denison, R.F. Measuring the fitness of symbiotic rhizobia. *Symbiosis* **2012**, *55*, 85–90. [CrossRef]
39. Denison, R.F.; Kiers, E.T. Lifestyle alternatives for rhizobia: Mutualism, parasitism, and forgoing symbiosis. *FEMS Microbiol. Lett.* **2004**, *237*, 187–193. [CrossRef]
40. Checcucci, A.; Azzarello, E.; Bazzicalupo, M.; Galardini, M.; Lagomarsino, A.; Mancuso, S.; Marti, L.; Marzano, M.C.; Mocali, S.; Squartini, A.; et al. Mixed nodule infection in Sinorhizobium meliloti-medicago sativa symbiosis suggest the presence of cheating behavior. *Front. Plant Sci.* **2016**, *7*, 835. [CrossRef]
41. Bellabarba, A.; Bacci, G.; Decorosi, F.; Aun, E.; Azzarello, E.; Remm, M.; Giovannetti, L. Competitiveness for Nodule Colonization in *Sinorhizobium meliloti*: Combined In Vitro-Tagged Strain Competition and Genome-Wide Association Analysis. *mSystems* **2021**, *6*, e00550-21. [CrossRef]
42. Gano-Cohen, K.A.; Wendlandt, C.E.; Stokes, P.J.; Blanton, M.A.; Quides, K.W.; Zomorrodian, A.; Adinata, E.S.; Sachs, J.L. Interspecific conflict and the evolution of ineffective rhizobia. *Ecol. Lett.* **2019**, *22*, 914–924. [CrossRef]
43. Sachs, J.L.; Simms, E.L. The origins of uncooperative rhizobia. *Oikos* **2008**, *117*, 961–966. [CrossRef]
44. Batstone, R.T.; Dutton, E.M.; Wang, D.; Yang, M.; Frederickson, M.E. The evolution of symbiont preference traits in the model legume *Medicago truncatula*. *New Phytol.* **2017**, *213*, 1850–1861. [CrossRef] [PubMed]
45. Heath, K.D.; Tiffin, P. Stabilizing mechanisms in a legume-rhizobium mutualism. *Evolution* **2009**, *63*, 652–662. [CrossRef] [PubMed]
46. Kiers, E.T.; Rousseau, R.A.; Denison, R.F. Measured sanctions: Legume hosts detect quantitative variation in rhizobium cooperation and punish accordingly. *Evol. Ecol.* **2006**, 1077–1086.
47. Burghardt, L.T.; Epstein, B.; Guhlin, J.; Nelson, M.S.; Taylor, M.R.; Young, N.D.; Sadowsky, M.J.; Tiffin, P. Select and resequence reveals relative fitness of bacteria in symbiotic and free-living environments. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 2425–2430. [CrossRef] [PubMed]
48. Pobigaylo, N.; Szymczak, S.; Nattkemper, T.W.; Becker, A. Identification of genes relevant to symbiosis and competitiveness in *Sinorhizobium meliloti* using signature-tagged mutants. *Mol. Plant-Microbe Interact.* **2008**, *21*, 219–231. [CrossRef] [PubMed]
49. Burghardt, L.T. Evolving together, evolving apart: Measuring the fitness of rhizobial bacteria in and out of symbiosis with leguminous plants. *New Phytol.* **2020**, *228*, 28–34. [CrossRef] [PubMed]
50. di Cenzo, G.C.; Zamani, M.; Milunovic, B.; Finan, T.M. Genomic resources for identification of the minimal N $_2$ -fixing symbiotic genome. *Environ. Microbiol.* **2016**, *18*, 2534–2547. [CrossRef] [PubMed]
51. Geddes, B.A.; Kearsley, J.V.S.; Huang, J.; Zamani, M.; Muhammed, Z.; Sather, L. Minimal gene set from *Sinorhizobium* (*Ensifer*) *meliloti* pSymA required for efficient symbiosis with *Medicago*. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2018015118. [CrossRef] [PubMed]
52. Checcucci, A.; Dicenzo, G.C.; Ghini, V.; Bazzicalupo, M.; Becker, A.; Decorosi, F.; Döhlemann, J.; Fagorzi, C.; Finan, T.M.; Fondi, M.; et al. Creation and Characterization of a Genomically Hybrid Strain in the Nitrogen-Fixing Symbiotic Bacterium *Sinorhizobium meliloti*. *ACS Synth. Biol.* **2018**, *7*, 2365–2378. [CrossRef] [PubMed]
53. Jones, K.M.; Kobayashi, H.; Davies, B.W.; Taga, M.E.; Walker, G.C. How rhizobial symbionts invade plants: The *Sinorhizobium—Medicago* model. *Nature* **2007**, *5*, 619–633. [CrossRef] [PubMed]
54. Burghardt, L.T.; Guhlin, J.; Chun, C.L.; Liu, J.; Sadowsky, M.J.; Stupar, R.M.; Young, N.D.; Tiffin, P. Transcriptomic basis of genome by genome variation in a legume-rhizobia mutualism. *Mol. Ecol.* **2017**, *26*, 6122–6135. [CrossRef]

55. Fagorzi, C.; Bacci, G.; Huang, R.; Cangioli, L.; Checcucci, A.; Fini, M.; Perrin, E.; Natali, C.; diCenzo, G.C.; Mengoni, A. Nonadditive transcriptomic signatures of genotype-by-genotype interactions during the initiation of plant-rhizobium symbiosis. *mSystems* **2021**, *6*, e00974-20. [CrossRef] [PubMed]
56. West, S.A.; Kiers, E.T.; Simms, E.L.; Denison, R.F. Sanctions and mutualism stability: Why do rhizobia fix nitrogen? *Proc. Biol. Sci.* **2002**, *269*, 685–694. [CrossRef]
57. Kiers, E.T.; Rousseau, R.A.; West, S.A.; Denison, R.F. Host sanctions and the legume–rhizobium mutualism. *Nature* **2003**, *425*, 78–81. [CrossRef]
58. Porter, S.S.; Simms, E.L. Selection for cheating across disparate environments in the legume-rhizobium mutualism. *Ecol. Lett.* **2014**, *17*, 1121–1129. [CrossRef] [PubMed]
59. Gubry-Rangin, C.; Garcia, M.; Béna, G. Partner choice in *Medicago truncatula-Sinorhizobium* symbiosis. *Proc. Biol. Sci.* **2010**, *277*, 1947–1951. [CrossRef] [PubMed]
60. Rogel, M.A.; Ormeno-Orrillo, E.; Martinez-Romero, E. Symbiovars in rhizobia reflect bacterial adaptation to legumes. *Appl. Microbiol.* **2011**, *34*, 96–104. [CrossRef] [PubMed]
61. Mergaert, P.; Van Montagu, M.; Holsters, M. Molecular mechanisms of Nod factor diversity. *Mol. Microbiol.* **1997**, *25*, 811–817. [CrossRef] [PubMed]
62. Wang, E.T.; Young, J.P.W. Symbiosis Genes: Organisation and Diversity. In *Ecology and Evolution of Rhizobia*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 123–144.
63. Andrews, M.; Andrews, M.E. Specificity in Legume-Rhizobia Symbioses. *Int. J. Mol. Sci.* **2017**, *18*, 705. [CrossRef] [PubMed]
64. de Souza, R.S.C.; Armanhi, J.S.L.; Damasceno, N.d.B.; Imperial, J.; Arruda, P. Genome Sequences of a Plant Beneficial Synthetic Bacterial Community Reveal Genetic Features for Successful Plant Colonization. *Front. Microbiol.* **2019**, *10*, 1779. [CrossRef]
65. Pini, F.; Galardini, M.; Bazzicalupo, M.; Mengoni, A. Plant-Bacteria Association and Symbiosis: Are There Common Genomic Traits in Alphaproteobacteria? *Genes* **2011**, *2*, 1017–1032. [CrossRef] [PubMed]
66. Cao, Y.; Halane, M.K.; Gassmann, W.; Stacey, G. The Role of Plant Innate Immunity in the Legume-Rhizobium Symbiosis. *Annu. Rev. Plant Biol* **2017**, *6812*, 535–561. [CrossRef] [PubMed]
67. Gourion, B.; Berrabah, F.; Ratet, P.; Stacey, G. Rhizobium-legume symbioses: The crucial role of plant immunity. *Trends Plant Sci.* **2015**, *20*, 186–194. [CrossRef] [PubMed]
68. Katalin, T.; Stacey, G. Does plant immunity play a critical role during initiation of the legume-rhizobium symbiosis? *Front. Plant Sci.* **2015**, *6*, 401. [CrossRef]
69. Nelson, M.S.; Sadowsky, M.J. Secretion systems and signal exchange between nitrogen-fixing rhizobia and legumes. *Front. Plant Sci.* **2015**, *6*, 491. [CrossRef] [PubMed]
70. Marczak, M.; Mazur, A.; Koper, P.; Żebracki, K.; Skorupska, A. Synthesis of rhizobial exopolysaccharides and their importance for symbiosis with legume plants. *Genes* **2017**, *8*, 360. [CrossRef] [PubMed]
71. Niehaus, K.; Kapp, D.; Pühler, A. Plant defence and delayed infection of alfalfa pseudonodules induced by an exopolysaccharide (EPS I)-deficient *Rhizobium meliloti* mutant. *Planta* **1993**, *190*, 415–425. [CrossRef]
72. Liu, Y.H.; Wang, E.T.; Jiao, Y.S.; Tian, C.F.; Wang, L.; Wang, Z.J.; Guan, J.J.; Singh, R.P.; Chen, W.X.; Chen, W.F. Symbiotic characteristics of *Bradyrhizobium diazoefficiens* USDA 110 mutants associated with shrubby sophora (*Sophora flavescens*) and soybean (*Glycine max*). *Microbiol. Res.* **2018**, *214*, 19–27. [CrossRef] [PubMed]
73. Salas, M.E.; Lozano, M.J.; López, J.L.; Draghi, W.O.; Serrania, J.; Torres Tejerizo, G.A.; Albicoro, F.J.; Nilsson, J.F.; Pistorio, M.; Del Papa, M.F.; et al. Specificity traits consistent with legume-rhizobia coevolution displayed by *Ensifer meliloti* rhizosphere colonization. *Environ. Microbiol.* **2017**, *19*, 3423–3438. [CrossRef] [PubMed]
74. Fagorzi, C.; Bacci, G.; Huang, R.; Cangioli, L.; Checcucci, A.; Fini, M.; Perrin, E.; Natali, C.; DiCenzo, G.C.; Mengoni, A. Dissecting transcriptomic signatures of genotype x genotype interactions during the initiation of plant-rhizobium symbiosis. *bioRxiv* **2020**. [CrossRef]
75. Benedict, A.B.; Ghosh, P.; Scott, S.M.; Griffits, J.S. A conserved rhizobial peptidase that interacts with host-derived symbiotic peptides. *Sci. Rep.* **2021**, *11*, 11779. [CrossRef] [PubMed]
76. Perret, X.; Staehelin, C. Molecular Basis of Symbiotic Promiscuity. *Microbiol. Mol. Biol. Rev.* **2000**, *64*, 180–201. [CrossRef] [PubMed]
77. Cangioli, L.; Checcucci, A.; Mengoni, A.; Fagorzi, C. Legume tasters: Symbiotic rhizobia host preference and smart inoculant formulations. *Biol. Commun.* **2021**, *66*, 47–54. [CrossRef]
78. diCenzo, G.C.; MacLean, A.M.; Milunovic, B.; Golding, G.B.; Finan, T.M. Examination of Prokaryotic Multipartite Genome Evolution through Experimental Genome Reduction. *PLoS Genet.* **2014**, *10*, e1004742. [CrossRef] [PubMed]
79. Galardini, M.; Pini, F.; Bazzicalupo, M.; Biondi, E.G.; Mengoni, A. Replicon-dependent bacterial genome evolution: The case of Sinorhizobium meliloti. *Genome Biol. Evol.* **2013**, *5*, 542–558. [CrossRef] [PubMed]
80. DiCenzo, G.C.; Checcucci, A.; Bazzicalupo, M.; Mengoni, A.; Viti, C.; Dziewit, L.; Finan, T.M.; Galardini, M.; Fondi, M. Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*. *Nat. Commun.* **2016**, *7*, 12219. [CrossRef]
81. Oono, R.; Anderson, C.G.; Denison, R.F. Failure to fix nitrogen by non-reproductive symbiotic rhizobia triggers host sanctions that reduce fitness of their reproductive clonemates. *Proc. R. Soc. B Biol. Sci.* **2011**, *278*, 2698–2703. [CrossRef] [PubMed]

82. Mergaert, P.; Nikovics, K.; Kelemen, Z.; Maunoury, N.; Vaubert, D.; Kondorosi, A.; Kondorosi, E. A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. *Plant Physiol.* **2003**, *132*, 161–173. [CrossRef] [PubMed]
83. Van de Velde, W.; Zehirov, G.; Szatmari, A.; Debreczeny, M.; Ishihara, H.; Kevei, Z.; Farkas, A.; Mikulass, K.; Nagy, A.; Tiricz, H. Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* **2010**, *327*, 1122–1126. [CrossRef] [PubMed]
84. Horváth, B.; Domonkos, Á.; Kereszt, A.; Szűcs, A.; Ábrahám, E.; Ayaydin, F.; Bóka, K.; Chen, Y.; Chen, R.; Murray, J.D. Loss of the nodule-specific cysteine rich peptide, NCR169, abolishes symbiotic nitrogen fixation in the *Medicago truncatula* dnf7 mutant. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 15232–15237. [CrossRef] [PubMed]
85. Kim, M.; Chen, Y.; Xi, J.; Waters, C.; Chen, R.; Wang, D. An antimicrobial peptide essential for bacterial survival in the nitrogen-fixing symbiosis. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 201500123. [CrossRef] [PubMed]
86. Downie, J.A.; Kondorosi, E. Why Should Nodule Cysteine-Rich (NCR) Peptides Be Absent From Nodules of Some Groups of Legumes but Essential for Symbiotic N-Fixation in Others? *Front. Agron.* **2021**, *3*, 654576. [CrossRef]
87. Nallu, S.; Silverstein, K.A.T.; Zhou, P.; Young, N.D.; VandenBosch, K.A. Patterns of divergence of a large family of nodule cysteine-rich peptides in accessions of *Medicago truncatula*. *Plant J.* **2014**, *78*, 697–705. [CrossRef] [PubMed]
88. Clay, K.; Kover, P.X. The Red Queen Hypothesis and Plant/Pathogen Interactions. *Annu. Rev. Phytopathol.* **1996**, *34*, 29–50. [CrossRef] [PubMed]
89. Price, P.A.; Tanner, H.R.; Dillon, B.A.; Shabab, M.; Walker, G.C.; Griffitts, J.S. Rhizobial peptidase HrrP cleaves host-encoded signaling peptides and mediates symbiotic compatibility. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 201417797. [CrossRef] [PubMed]
90. Checcucci, A.; Azzarello, E.; Bazzicalupo, M.; De Carlo, A.; Emiliani, G.; Mancuso, S.; Spini, G.; Viti, C.; Mengoni, A. Role and regulation of ACC deaminase gene in *Sinorhizobium meliloti:* Is it a symbiotic, rhizospheric or endophytic gene? *Front. Genet.* **2017**, *8*, 6. [CrossRef]
91. Okazaki, S.; Nukui, N.; Sugawara, M.; Minamisawa, K. Rhizobial Strategies to Enhance Symbiotic Interactions: Rhizobitoxine and 1-Aminocyclopropane-1-Carboxylate Deaminase. *Microbes Environ.* **2004**, *19*, 99–111. [CrossRef]
92. Ratcliff, W.C.; Denison, R.F. Rhizobitoxine producers gain more poly-3-hydroxybutyrate in symbiosis than do competing rhizobia, but reduce plant growth. *ISME J.* **2009**, *3*, 870–872. [CrossRef]
93. Ratcliff, W.C.; Kadam, S.V.; Denison, R.F. Poly-3-hydroxybutyrate (PHB) supports survival and reproduction in starving rhizobia. *FEMS Microbiol. Ecol.* **2008**, *65*, 391–399. [CrossRef] [PubMed]
94. Wang, C.; Saldanha, M.; Sheng, X.; Shelswell, K.J.; Walsh, K.T.; Sobral, B.W.S.; Charles, T.C. Roles of poly-3-hydroxybutyrate (PHB) and glycogen in symbiosis of *Sinorhizobium meliloti* with *Medicago* sp. *Microbiology* **2007**, *153*, 388–398. [CrossRef]
95. Jiao, Y.S.; Liu, Y.H.; Yan, H.; Wang, E.T.; Tian, C.F.; Chen, W.X.; Guo, B.L.; Chen, W.F. Rhizobial diversity and nodulation characteristics of the extremely promiscuous legume Sophora flavescens. *Mol. Plant-Microbe Interact.* **2015**, *28*, 1338–1352. [CrossRef]
96. Ratcliff, W.C.; Denison, R.F. Report Individual-Level Bet Hedging in the Bacterium *Sinorhizobium meliloti*. *CURBIO* **2010**, *20*, 1740–1744. [CrossRef]
97. Rolli, E.; Vergani, L.; Ghitti, E.; Patania, G.; Mapelli, F.; Borin, S. 'Cry-for-help' in contaminated soil: A dialogue among plants and soil microbiome to survive in hostile conditions. *Environ. Microbiol.* **2021**, *23*, 5690–5703. [CrossRef] [PubMed]
98. Schulz-Bohm, K.; Gerards, S.; Hundscheid, M.; Melenhorst, J.; De Boer, W.; Garbeva, P. Calling from distance: Attraction of soil bacteria by plant root volatiles. *ISME J.* **2018**, *12*, 1252–1262. [CrossRef] [PubMed]
99. de la Porte, A.; Schmidt, R.; Yergeau, É.; Constant, P. A Gaseous Milieu: Extending the Boundaries of the Rhizosphere. *Trends Microbiol.* **2020**, *28*, 536–542. [CrossRef]
100. Liu, H.; Brettell, L.E. Plant Defense by VOC-Induced Microbial Priming. *Trends Plant Sci.* **2019**, *24*, 187–189. [CrossRef] [PubMed]
101. Soto, M.J.; López-Lara, I.M.; Geiger, O.; Romero-Puertas, M.C.; van Dillewijn, P. Rhizobial Volatiles: Potential New Players in the Complex Interkingdom Signaling With Legumes. *Front. Plant Sci.* **2021**, *12*, 1219. [CrossRef] [PubMed]
102. Venturi, V.; Keel, C. Signaling in the Rhizosphere. *Trends Plant Sci.* **2016**, *21*, 187–198. [CrossRef] [PubMed]
103. Gao, M.; Chen, H.; Eberhard, A.; Gronquist, M.R.; Robinson, J.B.; Rolfe, B.G.; Bauer, W.D. sinI- and expR-Dependent Quorum Sensing in *Sinorhizobium meliloti*. *J. Bacteriol.* **2005**, *187*, 7931–7944. [CrossRef] [PubMed]
104. Sanchez-Contreras, M.; Bauer, W.D.; Gao, M.; Robinson, J.B.; Downie, J.A. Quorum-sensing regulation in rhizobia and its role in symbiotic interactions with legumes. *Philos. Trans. R. Soc. B Biol. Sci.* **2007**, *362*, 1149–1163. [CrossRef] [PubMed]
105. Gao, M.; Teplitski, M.; Robinson, J.B.; Bauer, W.D. Production of substances by *Medicago truncatula* that affect bacterial quorum sensing. *Mol. Plant-Microbe Interact.* **2003**, *16*, 827–834. [CrossRef] [PubMed]
106. Martínez-Romero, E.; Aguirre-Noyola, J.L.; Taco-Taype, N.; Martínez-Romero, J.; Zuñiga-Dávila, D. Plant microbiota modified by plant domestication. *Syst. Appl. Microbiol.* **2020**, *43*, 126106. [CrossRef]
107. Chaparro, J.M.; Badri, D.V.; Vivanco, J.M. Rhizosphere microbiome assemblage is affected by plant development. *ISME J.* **2014**, *8*, 790–803. [CrossRef]
108. Schütz, V.; Frindte, K.; Cui, J.; Zhang, P.; Hacquard, S.; Schulze-Lefert, P.; Knief, C.; Schulz, M.; Dörmann, P. Differential Impact of Plant Secondary Metabolites on the Soil Microbiota. *Front. Microbiol.* **2021**, *12*, 1267. [CrossRef] [PubMed]
109. Knee, E.M.; Gong, F.-C.; Gao, M.; Teplitski, M.; Jones, A.R.; Foxworthy, A.; Mort, A.J.; Bauer, W.D. Root mucilage from pea and its utilization by rhizosphere bacteria as a sole carbon source. *Mol. Plant-Microbe Interact.* **2001**, *14*, 775–784. [CrossRef] [PubMed]

110. Moody, S.F.; Clarke, A.E.; Bacic, A. Structural analysis of secreted slime from wheat and cowpea roots. *Phytochemistry* **1988**, *27*, 2857–2861. [CrossRef]
111. Bringhurst, R.M.; Cardon, Z.G.; Gage, D.J. Galactosides in the rhizosphere: Utilization by *Sinorhizobium meliloti* and development of a biosensor. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4540–4545. [CrossRef]
112. Liu, Y.; Jiang, X.; Guan, D.; Zhou, W.; Ma, M.; Zhao, B.; Cao, F.; Li, L.; Li, J. Transcriptional analysis of genes involved in competitive nodulation in *Bradyrhizobium diazoefficiens* at the presence of soybean root exudates. *Sci. Rep.* **2017**, *7*, 10946. [CrossRef] [PubMed]
113. Klonowska, A.; Melkonian, R.; Miché, L.; Tisseyre, P.; Moulin, L. Transcriptomic profiling of *Burkholderia phymatum* STM815, *Cupriavidus taiwanensis* LMG19424 and *Rhizobium mesoamericanum* STM3625 in response to *Mimosa pudica* root exudates illuminates the molecular basis of their nodulation competitiveness and symbiotic ev. *BMC Genomics* **2018**, *19*, 105. [CrossRef] [PubMed]
114. Li, B.; Li, Y.-Y.; Wu, H.-M.; Zhang, F.-F.; Li, C.-J.; Li, X.-X.; Lambers, H.; Li, L. Root exudates drive interspecific facilitation by enhancing nodulation and N2 fixation. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 6496–6501. [CrossRef] [PubMed]
115. Liu, C.W.; Murray, J.D. The role of flavonoids in nodulation host-range specificity: An update. *Plants* **2016**, *5*, 33. [CrossRef] [PubMed]
116. Peck, M.C.; Fisher, R.F.; Long, S.R. Diverse Flavonoids Stimulate NodD1 Binding to nod Gene Promoters in *Sinorhizobium meliloti*. *J. Bacteriol.* **2006**, *188*, 5417–5427. [CrossRef] [PubMed]
117. Spini, G.; Decorosi, F.; Cerboneschi, M.; Tegli, S.; Mengoni, A.; Viti, C.; Giovannetti, L. Effect of the plant flavonoid luteolin on Ensifer meliloti 3001 phenotypic responses. *Plant Soil* **2016**, *399*, 159–178. [CrossRef]
118. Compton, K.K.; Hildreth, S.B.; Helm, R.F.; Scharf, B.E. An Updated Perspective on *Sinorhizobium meliloti* Chemotaxis to Alfalfa Flavonoids. *Front. Microbiol.* **2020**, *11*, 581482. [CrossRef]
119. Janczarek, M.; Rachwał, K.; Marzec, A.; Grzadziel, J.; Palusińska-Szysz, M. Signal molecules and cell-surface components involved in early stages of the legume-rhizobium interactions. *Appl. Soil Ecol.* **2014**, *85*, 94–113. [CrossRef]
120. Di Cenzo, G.C.; Finan, T.M. The Divided Bacterial Genome: Structure, Function, and Evolution. *Microbiol. Mol. Biol. Rev.* **2017**, *81*, e00019-17.
121. Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, F.; Barnett, M.J.; Becker, A.; Boistard, P.; et al. The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **2001**, *293*, 668–672. [CrossRef]
122. Santamarı, R.I.; Medrano-soto, A.; Bustos, P.; Herna, I.; Moreno-hagelsieb, G.; Janga, S.C.; Ramı, M.A.; Collado-vides, J.; Da, G. The partitioned *Rhizobium etli* genome: Genetic and metabolic redundancy in seven interacting replicons. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 3834–3839.
123. Young, J.P.W.; Crossman, L.C.; Johnston, A.W.B.; Thomson, N.R.; Ghazoui, Z.F.; Hull, K.H.; Wexler, M.; Curson, A.R.J.; Todd, J.D.; Poole, P.S.; et al. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Young* **2006**, *7*, R34. [CrossRef]
124. Fagorzi, C.; Ilie, A.; Decorosi, F.; Cangioli, L.; Viti, C.; Mengoni, A.; diCenzo, G.C. Symbiotic and non-symbiotic members of the genus *Ensifer* (syn. *Sinorhizobium*) are separated into two clades based on comparative genomics and high-throughput phenotyping. *Genome Biol. Evol.* **2020**, *12*, 2521–2534. [CrossRef] [PubMed]
125. Ochman, H.; Moran, N.A. Genes lost and genes found: Evolution of bacterial pathogenesis and symbiosis. *Science* **2001**, *292*, 1096–1099. [CrossRef]
126. Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Gurjal, M.; Hong, A.; Huizar, L.; Bowser, L.; Capela, D.; et al. Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9883–9888. [CrossRef] [PubMed]
127. Oresnik, I.J.; Liu, S.-L.; Yost, C.K.; Hynes, M.F. Megaplasmid pRme2011a of *Sinorhizobium meliloti* is not required for viability. *J. Bacteriol.* **2000**, *182*, 3582–3586. [CrossRef]
128. Ma, W.; Charles, T.C.; Glick, B.R. Expression of an exogenous 1-aminocyclopropane-1-carboxylate deaminase gene in *Sinorhizobium meliloti* increases its ability to nodulate alfalfa. *Appl. Environ. Microbiol.* **2004**, *70*, 5891–5897. [CrossRef] [PubMed]
129. Pini, F.; Spini, G.; Galardini, M.; Bazzicalupo, M.; Benedetti, A.; Chiancianesi, M.; Florio, A.; Lagomarsino, A.; Migliore, M.; Mocali, S.; et al. Molecular phylogeny of the nickel-resistance gene *nreB* and functional role in the nickel sensitive symbiotic nitrogen fixing bacterium *Sinorhizobium meliloti*. *Plant Soil* **2013**, *377*, 189–201. [CrossRef]
130. Galardini, M.; Brilli, M.; Spini, G.; Rossi, M.; Roncaglia, B.; Bani, A.; Chiancianesi, M.; Moretto, M.; Engelen, K.; Bacci, G.; et al. Evolution of Intra-specific Regulatory Networks in a Multipartite Bacterial Genome. *PLoS Comput. Biol.* **2015**, *11*, e1004478. [CrossRef] [PubMed]
131. diCenzo, G.C.; Benedict, A.B.; Fondi, M.; Walker, G.C.; Finan, T.M.; Mengoni, A.; Griffitts, J.S. Robustness encoded across essential and accessory replicons of the ecologically versatile bacterium *Sinorhizobium meliloti*. *PLoS Genet.* **2018**, *14*, e1007357. [CrossRef] [PubMed]
132. Zheng, H.; Mao, Y.; Zhu, Q.; Ling, J.; Zhang, N.; Naseer, N.; Zhong, Z.; Zhu, J. The quorum sensing regulator CinR hierarchically regulates two other quorum sensing pathways in ligand-dependent and-independent fashions in *Rhizobium etli*. *J. Bacteriol.* **2015**, *197*, 1573–1581. [CrossRef]
133. Danino, V.E.; Wilkinson, A.; Edwards, A.; Downie, J.A. Recipient-induced transfer of the symbiotic plasmid pRL1JI in *Rhizobium leguminosarum* bv. *viciae* is regulated by a quorum-sensing relay. *Mol. Microbiol.* **2003**, *50*, 511–525. [PubMed]

134. Zhang, L.H.; Kerr, A. A diffusible compound can enhance conjugal transfer of the Ti plasmid in *Agrobacterium tumefaciens*. *J. Bacteriol.* **1991**, *173*, 1867–1872. [CrossRef] [PubMed]
135. Blanca-Ordóñez, H.; Oliva-García, J.J.; Pérez-Mendoza, D.; Soto, M.J.; Olivares, J.; Sanjuán, J.; Nogales, J. pSymA-dependent mobilization of the *Sinorhizobium meliloti* pSymB megaplasmid. *J. Bacteriol.* **2010**, *192*, 6309–6312. [CrossRef] [PubMed]
136. Pistorio, M.; Giusti, A.; Papa, F.D.; Draghi, W.O.; Lozano, M.J.; Tejerizo, T.; Lagares, A. Conjugal properties of the *Sinorhizobium meliloti* plasmid mobilome. *FEMS Microbiol. Ecol.* **2008**, *65*, 372–382. [CrossRef] [PubMed]
137. Nogales, J.; Blanca-Ordóñez, H.; Olivares, J.; Sanjuán, J. Conjugal transfer of the *Sinorhizobium meliloti* 1021 symbiotic plasmid is governed through the concerted action of one- and two-component signal transduction regulators. *Environ. Microbiol.* **2013**, *15*, 811–821. [CrossRef]
138. Pérez-Mendoza, D.; Sepúlveda, E.; Pando, V.; Munoz, S.; Nogales, J.; Olivares, J.; Soto, M.J.; Herrera-Cervera, J.A.; Romero, D.; Brom, S. Identification of the *rctA* gene, which is required for repression of conjugative transfer of rhizobial symbiotic megaplasmids. *J. Bacteriol.* **2005**, *187*, 7341–7350. [CrossRef]
139. Wozniak, R.A.F.; Waldor, M.K. Integrative and conjugative elements: Mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat. Rev. Microbiol.* **2010**, *8*, 552–563. [CrossRef] [PubMed]
140. Ling, J.; Wang, H.; Wu, P.; Li, T.; Tang, Y.; Naseer, N.; Zheng, H.; Masson-Boivin, C. Plant nodulation inducers enhance horizontal gene transfer of *Azorhizobium caulinodans* symbiosis island. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 13875–13880. [CrossRef] [PubMed]
141. Hassani, M.A.; Durán, P.; Hacquard, S. Microbial interactions within the plant holobiont. *Microbiome* **2018**, *6*, 58. [CrossRef]
142. Duran, P.; Thiergart, T.; Garrido-Oter, R.; Agler, M.; Kemen, E.; Schulze-Lefert, P.; Hacquard, S. Microbial interkingdom interactions in roots promote *Arabidopsis* survival. *Cell* **2018**, *175*, 973–983. [CrossRef] [PubMed]
143. Checcucci, A.; Marchetti, M. The Rhizosphere Talk Show: The Rhizobia on Stage. *Front. Agron.* **2020**, *2*, 591494. [CrossRef]
144. Bonfante, P.; Genre, A. Mechanisms underlying beneficial plant—Fungus interactions in mycorrhizal symbiosis. *Nat. Commun.* **2010**, *1*, 48. [CrossRef] [PubMed]
145. Bozsoki, Z.; Cheng, J.; Feng, F.; Gysel, K.; Vinther, M.; Andersen, K.R.; Oldroyd, G.; Blaise, M.; Radutoiu, S.; Stougaard, J. Receptor-mediated chitin perception in legume roots is functionally separable from Nod factor perception. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E8118–E8127. [CrossRef] [PubMed]
146. Antunes, P.M.; Rajcan, I.; Goss, M.J. Specific flavonoids as interconnecting signals in the tripartite symbiosis formed by arbuscular mycorrhizal fungi, *Bradyrhizobium japonicum* (Kirchner) Jordan and soybean (*Glycine max* (L.) Merr.). *Soil Biol. Biochem.* **2006**, *38*, 533–543. [CrossRef]
147. Afkhami, M.E.; Friesen, M.L.; Stinchcombe, J.R. Multiple Mutualism Effects generate synergistic selection and strengthen fitness alignment in the interaction between legumes, rhizobia and mycorrhizal fungi. *Ecol. Lett.* **2021**, *24*, 1824–1834. [CrossRef]
148. Kaschuk, G.; Leffelaar, P.A.; Giller, K.E.; Alberton, O.; Hungria, M.; Kuyper, T.W. Responses of legumes to rhizobia and arbuscular mycorrhizal fungi: A meta-analysis of potential photosynthate limitation of symbioses. *Soil Biol. Biochem.* **2010**, *42*, 125–127. [CrossRef]
149. Horiuchi, J.I.; Prithiviraj, B.; Bais, H.P.; Kimball, B.A.; Vivanco, J.M. Soil nematodes mediate positive interactions between legume plants and rhizobium bacteria. *Planta* **2005**, *222*, 848–857. [CrossRef] [PubMed]
150. Bellabarba, A.; Fagorzi, C.; George, C.; Pini, F.; Viti, C.; Checcucci, A. Deciphering the Symbiotic Plant Microbiome: Translating the Most Recent Discoveries on Rhizobia for the Improvement of Agricultural Practices in Metal-Contaminated and High Saline Lands. *Agronomy* **2019**, *9*, 529. [CrossRef]
151. Schulte, C.C.M.; Ramachandran, V.K.; Papachristodoulou, A.; Poole, P.S. Genome-scale metabolic modelling of lifestyle changes in Rhizobium leguminosarum. *bioRxiv* **2021**, *85*, 693–702. [CrossRef] [PubMed]
152. Schulte, C.C.M.; Borah, K.; Wheatley, R.M.; Terpolilli, J.J.; Saalbach, G.; Crang, N.; Groot, D.H.D.; Ratcliffe, R.G.; Kruger, N.J.; Papachristodoulou, A.; et al. Metabolic control of nitrogen fixation in rhizobium-legume symbioses. *Sci. Adv.* **2021**, *7*, eabh2433. [CrossRef] [PubMed]
153. Klein, M.; Stewart, J.D.; Porter, S.S.; Weedon, J.T.; Kiers, E.T. The evolution of manipulative microbial behaviors in the rhizosphere. *Evol. Appl.* **2021**, 1–16. [CrossRef]
154. Bevivino, A.; Bacci, G.; Drevinek, P.; Nelson, M.T.; Hoffman, L.; Mengoni, A. Deciphering the Ecology of Cystic Fibrosis Bacterial Communities: Towards Systems-Level Integration. *Trends Mol. Med.* **2019**, *25*, 1110–1122. [CrossRef]

SYMBIOGENETICS

Legume tasters: symbiotic rhizobia host preference and smart inoculant formulations

Lisa Cangioli[1], Alice Checcucci[2], Alessio Mengoni[1], and Camilla Fagorzi[1]

[1]Department of Biology, University of Florence,
Via Madonna del Piano, 6, Sesto Fiorentino, 50019, Italy
[2]Department of Agricultural and Food Science, University of Bologna,
Viale Giuseppe Fanin, 40-50, Bologna, 40127, Italy

Address correspondence and requests for materials to Camilla Fagorzi,
camilla.fagorzi@unifi.it

Abstract

Mutualistic interactions have great importance in ecology, with genetic information that takes shape through interactions within the symbiotic partners and between the partners and the environment. It is known that variation of the host-associated microbiome contributes to buffer adaptation challenges of the host's physiology when facing varying environmental conditions. In agriculture, pivotal examples are symbiotic nitrogen-fixing rhizobia, known to contribute greatly to host (legume plants) adaptation and host productivity. A holistic view of increasing crop yield and resistance to biotic and abiotic stresses is that of microbiome engineering, the exploitation of a host-associated microbiome through its rationally designed manipulation with synthetic microbial communities. However, several studies highlighted that the expression of the desired phenotype in the host resides in species-specific, even genotype-specific interactions between the symbiotic partners. Consequently, there is a need to dissect such an intimate level of interaction, aiming to identify the main genetic components in both partners playing a role in symbiotic differences/host preferences. In the present paper, while briefly reviewing the knowledge and the challenges in plant–microbe interaction and rhizobial studies, we aim to promote research on genotype x genotype interaction between rhizobia and host plants for a rational design of synthetic symbiotic nitrogen-fixing microbial communities to be used for sustainably improving leguminous plants yield.

Keyword: microbiome engineering, genome x genome, legume-rhizobia mutualism, symbiosis, precision agriculture

Nobody is an island: the relevance of mutualistic interactions with eukaryotic hosts

Mutualistic interactions occur everywhere in the biosphere and are pivotal in evolution, as well as having great importance in ecology (Bronstein, 2015). Mutualism involves co-evolution, innovation, change and the involvement of partners with complementary skills (Lanier et al., 2017). In 2008, the hologenome theory of evolution defined the holobiont as a unit of selection in evolution (Zilber-Rosenberg and Rosenberg, 2008). The eukaryotic host and its associated microbes are considered a superorganism, in which genetic information takes shape through interactions between the partners and between the partners and the environment (Wilson and Sober, 1989; Zilber-Rosenberg and Rosenberg, 2008). For example, plants and animals have extended their metabolic repertoire through the establishment of specific microbial communities associated to roots and gut, respectively. These microbial communities are characterized by the presence of bacteria, archaea, fungi, oomycetes, as well as viruses, and they can be partly considered as the host's extended genome, improving the nutrient acquisition process

Citation: Cangioli, L., Checcucci, A., Mengoni, A., and Fagorzi, C. 2021. Legume tasters: symbiotic rhizobia host preference and smart inoculant formulations. Bio. Comm. 66(1): 47–54. https://doi.org/10.21638/spbu03.2021.106

Authors' information: Lisa Cangioli, PhD Student, orcid.org/0000-0002-7714-1704; Alice Checcucci, PhD, Research Fellow, orcid.org/0000-0002-0019-0997; Alessio Mengoni, PhD, Associate Professor, Head of Laboratory, orcid.org/0000-0002-1265-8251; Camilla Fagorzi, PhD Student, orcid.org/0000-0002-9996-8868

Manuscript Editor: Anton Nizhnikov, Department of Genetics and Biotechnology, Faculty of Biology, Saint Petersburg State University, Saint Petersburg, Russia

Received: October 19, 2020;

Revised: December 21, 2020;

Accepted: January 3, 2021.

Copyright: © 2021 Cangioli et al. This is an open-access article distributed under the terms of the License Agreement with Saint Petersburg State University, which permits to the authors unrestricted distribution, and self-archiving free of charge.

Funding: This work was supported by the grant MICRO4Legumes ("Il microbioma vegetale simbionte come strumento per il miglioramento delle leguminose foraggere. Acronimo"), D.M.n.89267 (Italian Ministry of Agriculture).

Competing interests: The authors have declared that no competing interests exist.

in the animal gut and in plant roots (Turner et al., 2013; Hacquard et al., 2015). However, how this can imply co-evolution and consideration of the holobiont as a unit of selection is still hotly debated (Moran and Sloan, 2015; Koskella and Bergelson, 2020). Indeed, the host-associated microbiota is not fully stable, and rapid variation in its composition and functioning can take place. Such variations may contribute to buffer adaptation challenges of the host's physiology when facing varying environmental conditions. In other words, the holobiont can allow more time-effective adaption in rapidly changing environments, reducing the actual selective pressure on the host genome (Zilber-Rosenberg and Rosenberg, 2008; Rosenberg and Zilber-Rosenberg, 2016).

Interactions between the host and microbes are shaped by multiple factors associated with host immunity, phylogeny, environmental features (e.g., pH, presence of biotic or abiotic compounds, temperature), diet and nutrient availability (Brinker et al., 2019).

Classically, plant domestication has involved selection of varieties on the basis of phenotypic features relevant to the farmer and adapted to the agricultural practices, without considering the associated microbiota. Since crops mostly rely on external inputs (e.g., fertilizers) and not on associated microbes, a decrease of microbiota diversity has been observed (Escudero-Martinez and Bulgarelli, 2019; Martínez-Romero et al., 2020), reducing the buffering effect offered by a diverse holobiont. For instance, in legumes it has been demonstrated that in the wild-growing varieties the symbiotic potential is usually higher than in commercial ones (Provorov and Tikhonovich, 2003).

Good diversity at the holobiont level may then be crucial in the search for environmentally sustainable crop production. Here, the authors explored the knowledge and challenges in the promotion of a rational use of symbiotic nitrogen-fixing bacteria for leguminous plant growth.

Biotechnology on the holobiont: cracking the microbiome

The increase of the human world population in the last years (currently 7 billion people, with 9 billion people expected by 2050) has led to the need to increase agricultural production (FAO; 2018). Inorganic nitrogen fertilizers are required to increase field production, but the excessive use of these compounds has deleterious effects on the environment (Yang and Fang, 2015; Zheng et al., 2019).

One of the possible alternative and sustainable approaches to enrich fields with nitrogen is the inoculation of crops with microorganisms able to promote the plant growth and health, the so-called *Plant Growth Promoting Microorganisms* (PGPM) (Lucy et al., 2004; Schlaeppi and Bulgarelli, 2015). In particular, *Plant Growth Promoting Rhizobacteria* (PGPR) can be applied as biocontrol agents, bio-inoculants and bio-fertilizers (Bloemberg and Lugtenberg, 2001). PGPM and PGPR constitute a relevant fraction of the plant mutualistic microbiome. They thrive on and within plants, using plant-produced organic molecules for their growth and "rewarding" the host with increased nutrient availability for the root apparatus as well as additional phytohormones (Werner et al., 2014).

Modification of the plant-associated microbiome (mainly, but not only, involving PGPR) has an effective potential to improve plant yield and resistance to biotic and abiotic stresses; this potential is stirring the attention of many investigators (Turner et al., 2013). This approach of "*microbiome engineering*" is based on reconstructing synthetic microbial communities after laboratory selection of microbes with the best ability to deliver PGP traits to the plants (Ke et al., 2020). The principal aims of microbiome manipulation are to: 1) reduce the incidence of plant diseases (Malfanova, 2013) ; 2) increase agricultural production (Bakker et al., 2012); 3) reduce chemical inputs (Adesemoye et al., 2009) and 4) reduce emissions of greenhouse gases (Singh et al., 2010). These goals meet the principles of agronomic sustainability and the world's increasing population (Turner et al., 2013). To date, successful attempts have been reviewed by de Vries et al., 2020; Dubey et al., 2020; Qiu et al., 2019; and Sudheer et al., 2020.

Legume–rhizobia mutualism as a model

Among the "rewards" from microbes, assimilable nitrogen compounds are one of the key components: the supply of nitrogen in the soil represents the classical limiting factors of plant productivity (Fageria and Baligar, 2005). Bacteria called rhizobia are the typical example of PGPM which provides the host plant assimilable nitrogen. Through a process called biological nitrogen fixation (BNF), involving several genes and signaling molecules, rhizobia fix atmospheric nitrogen into compounds assimilable by plants. The BNF is a paradigmatic example of mutualistic association between plants and bacteria. The symbiotic BNF (SNF, *Symbiotic Nitrogen Fixation*) is a facultative symbiotic association involving rhizobia (bacteria from Alfa- and Betaproteobacteria classes) and some actinobacteria (*Frankia* spp.), which associate with plants from the Fabales (*Leguminosae*), Fagales, Cucurbitales and Rosales orders (Pawlowski and Newton, 2008; Griesmann et al., 2018). During the symbiotic interaction, bacteria induce the formation of a specific structure at the root level, the nodule, then colonize nodule plant cells intracellularly and activate the key component of BNF (i.e., the nitrogenase enzyme), producing ammonia, which is then exchanged for the

photosynthetic products with the host plant (Kereszt et al., 2011).

The symbiosis between legumes and rhizobia is the most diffused and relevant in providing fixed nitrogen to agroecosystems. This interaction is mediated by a plethora of molecular signals that direct bacterial invasion, modulate the host defence response, permit the intracellular colonization, guide the regulation of the cell cycle and regulate the nutrient exchange (Jones et al., 2007; Gibson et al., 2008). One of the most essential signalling compounds produced by rhizobia is the Nod Factor (NF). Rhizobia are able to produce Nod factors characterized by different structures which, consequently, are recognized by specific host plant receptors (Geurts and Bisseling, 2002). Due to this process, evidence has emerged for coevolution between rhizobia and host plant at the population level (Igolkina et al., 2019). However, restricted host specificity is present, and phylogenetical studies confirmed that the limited host range evolved from an ancestral broad range mutualism (Pueppke and Broughton, 1999). Therefore, evolution has shaped the two partners toward species-specific and even strain-specific interactions allowing plant selection for the best beneficial symbionts (Checcucci et al., 2016; Remigi et al., 2016; Westhoek et al., 2017; Sachs et al., 2018).

The two-way exchange of benefits between plant roots and rhizobia offers us the basic knowledge to mold and manipulate this symbiosis interaction, sustaining the ecological and agronomic practices in agricultural systems (Bakker et al., 2012; Checcucci et al., 2017; Soumare et al., 2020).

Forming the symbiotic structure: recognition of the good partners

The multi-step process which allows the symbiotic interaction between nitrogen-fixing rhizobia and leguminous plants is highly regulated, to allow the selection by the plant of the most effective (highly rewarding) partners (Sachs et al., 2018). Many genes and mechanisms involved in the establishment of the symbiosis between rhizobia and host legumes have been identified (Roy et al., 2020).

The early events start with the exchange of signals between the nitrogen-fixing rhizobia and the plant in the rhizosphere. Here, the first specific signals are found in compounds present in root exudates (flavonoids), which bind receptor proteins in the rhizobial cell (NodD). Activated NodD trigger bacterial *nod* genes expression which lead to synthesis of the Nod factor (a lipochitooligosaccharide molecule), which is then secreted by the bacterial cell. The perception of the Nod factor by the plant roots induces developmental changes such as root hair curling, membrane depolarization, intracellular calcium oscillations, and the initiation of cell division in the root cortex, which establishes a meristem and nodule primordium (Liu et al., 2020). Species-specific early recognition is based on the fact that different rhizobial species and strains can have differential activation by flavonoids on NodD proteins and may produce different Nod factor molecules. Different Nod factors are then specifically recognized by plant receptors (LysM), ensuring species-specific recognition on the plant side (Bozsoki et al., 2020).

When the dialogue between plant roots and rhizobia is successful, a small number of bacteria are trapped by the root hair, which begins an inverse tip growth, forming a long and narrow passage called the "infection thread". The bacteria reach the root cortex as the final destination and reacquire properties of stem cells, which form the lateral organ nodule. The bacteria infect the nodule, enter the cytoplasm of plant cells and differentiate into a distinct cell type called bacteroid, which can fix atmospheric nitrogen into ammonia, establishing an intricate metabolic interchange with the host plant (diCenzo et al., 2020; Roy et al., 2020). Moreover here, nodules where rhizobia fail to produce fixed nitrogen are sanctioned, ensuring that only good mutualist rhizobia are allowed to thrive in root nodules (Kiers and Denison, 2008), though cheating may occur (Checcucci et al., 2016). Though molecular details of sanctions are still elusive, molecular determinants of species-specificity have been discovered. For instance, in the model legume *Medicago truncatula*, the plant response after bacteria invasion of the host cell is modulated by two nodule-specific cysteine-rich (NCR) peptides encoded by the genes *NFS1* and *NFS2* (Wang et al., 2018; Tirichine et al., 2000). More recently, again in *M. truncatula*, MtNPD1, a nodule-specific polycistin lipogenase, has been shown to represent an important determinant of host-strain specificity, with particular regard to plant invasion and nitrogen fixation (Pislariu et al., 2019). Functional symbiotic compatibility between *Mesorhizobium loti* and *Lotus japonicum* is determined by the presence of the nepenthesin-type aspartic peptidase nodule-induced LjAPN1 (Yamaya-Ito et al., 2018).

In *Pisum sativum*, Nod factors perception is mediated by the gene PsSym2, which probably encodes a LysM-RLK. PsSym2 was the first discovered symbiosis-related pea gene and it determines increased selectivity toward rhizobia in the pea cv. Afghanistan (Sulima et al., 2017).

The role of genotype x genotype interactions in shaping the symbiotic partnership

Interactions between different species with different genotypes imply the expression of a phenotype that depends on the interacting organisms (Bose and Schulte, 2014). The effects of (microbial) genotype (G) x (host) genotype (G) interactions can be relevant for both rate and direction of evolutionary selection (Wade, 2007) and have clear importance for rational improvement of plant–rhizobium

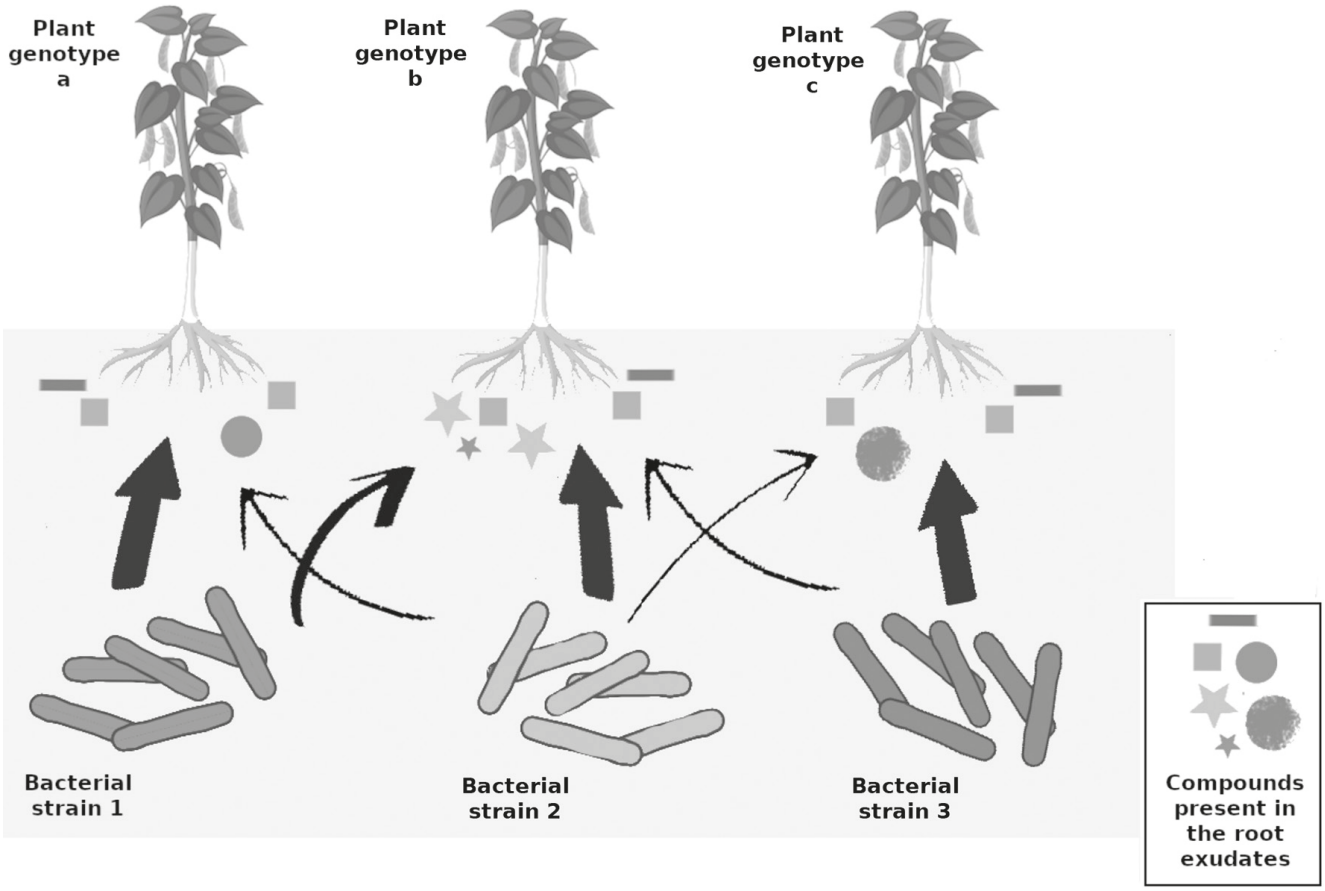


Fig. 1. Simplified model of genotype x genotype rhizobia–legume interaction. Different genotypes of the same species of legume interact with different genotypes of the same rhizobial species. Interactions can be trained by the expression of an intricate network of genes of both symbiotic partners, driving host preference. The rhizobial strain can initiate a mutualistic relationship with multiple legume genotypes, but specific genotype x genotype interaction can result in a more effective symbiotic ability. In the figure, the effectiveness of symbiotic ability is represented by arrows of different thickness (arrows are thicker when symbiosis is more effective) and compounds produced by the different plant genotypes are represented by the symbols near the root system (symbols of the same colour of the bacterial strain are responsible for a more effective symbiotic relationship).

symbiosis. As reviewed above, the development and application of genetic and genomic tools has made it possible to dissect in great detail the molecular determinants involved in plant–microbe interaction (Levy et al., 2018) and rhizobial symbioses (diCenzo et al., 2019).

However, most of the studies involved model strains and plants and their respective mutants, with only a limited evaluation of the genetic bases of the somewhat large phenotypic variation in the symbiotic phenotype due to natural genotype variation. Early studies in the alfalfa rhizobial symbiont *Sinorhizobium meliloti* showed that plant variety exerts a detectable effect on the recruitment of native soil rhizobial population (Carelli et al., 2000; Paffetti et al., s.d.). Clearly, the fitness advantage of the symbiosis (Burghardt, 2020) is influenced by GxG interactions between the two (or more) partners (Heath, 2008). Recently, the use a "select and resequence" approach has disclosed some of the determinants of GxG interaction in the model symbiosis between strains of *Sinorhizobium meliloti* and genotypes of the legume the *Medicago truncatula* (Burghardt et al., 2018, 2020). With a similar aim, a recent study has been conducted to analyse at the gene expression level the genotype-based variation of the interaction, using *S. meliloti* and the forage legume *M. sativa* (Kang et al., 2020). In this last work, the transcriptional profile of alfalfa inoculated with different *S. meliloti* strains was investigated, finding that the plant differentially expressed genes depending on the *S. meliloti* strain tested. Among the candidate genes influenced by the rhizobial strain genotype, genes encoding nodulins, NCR peptides and proteins belonging to the NBS-LRR family were identified (Kang et al., 2020). On the rhizobial side we recently sequenced the transcriptome of several *S. meliloti* strains incubated in the presence of root exudates of different alfalfa varieties. Results showed an intricate network of genes whose expression is influenced by the genotype of both symbiotic partners, suggesting the presence of a large number of determinants involved in GxG interactions which can have an impact on the outcome of the rhizobium–legume symbiosis (Fig. 1) (Fagorzi et al., 2021). Dissection of such a complex network, iden-

tifying the main components playing a role in symbiotic differences/host preferences, will make it possible to rationally screen and genetically improve the most promising plant and symbiotic bacteria partnerships, as well as to better understand the intricate routes of the evolution of symbioses.

On the plant side, the GxG interactions between selected nitrogen-fixing rhizobia and selected symbiotically active plant genotypes also lend advantages in terms of plant productivity (Provorov & Tikhonovich, 2003).

Holobiont improvement programs toward precision agriculture

Classical studies, such as the ones performed by Bliss (1990) and Barnes et al. (1984), showed the importance of a multidisciplinary approach to improve the symbiotic activity in *Phaseolus vulgaris* and *Medicago sativa*.

The study of plant genotype and breeding programs have been fundamental in understanding plant physiological traits and find directions for productivity improvement. The use of fertilizers can definitely be reduced by exploiting the knowledge of a plant's genome, soil conditions and plant-associated microorganisms. Indeed, genome editing methods allow targeted modification in the plant genotype, making it possible to improve and adapt a crop's genetic traits with the current soil characteristics (Joseph et al., 2020), including rhizosphere resident bacteria and fungi. The advantages of this technological and biological knowledge in the agricultural field were recently discussed in two workshops organized by the USDA National Institute of Food and Agriculture (NIFA)-funded Big Data Driven Agriculture, where the importance of their applicability was assessed on the basis of standardized protocol applications, funding opportunities and involvement of a broader scientific community and farmers (Shakoor et al., 2019). Currently, high-throughput techniques for plants and soil phenotyping, tools for the measurement of crop productivity, genome editing methods, breeding and microbial inoculation technologies form the basis of precision agriculture, which involves both soil/plant science experts and farm management operators.

For SNF, besides the biotic factors shaping legume–rhizobia nitrogen fixation (e.g., selection and interaction with the partner, competition with indigenous microorganisms), abiotic variables can strongly influence the success of the symbiosis. Indeed, nutrient availability, extreme temperature, pH, salinity and drought (Bellabarba et al., 2019) challenge legume crops, stimulating the need for an adjustable agriculture. Precision agriculture can ensure the health and productivity of soil and crops through knowledge and technology advancements, guaranteeing protection and well-being of the environment.

Presently, several diagnostic methods are based on the quantification of nitrogen in legumes, evaluating the effectiveness of rhizobial symbiosis (e.g., measures of nitrogen at the leaf level and measurement of nitrogen abundance in plant tissues through isotope-based methodologies (Thilakarathna and Raizada, 2018). Furthermore, several technologies aimed to optimize biological nitrogen fixation in the field are based on the diagnosis of soil traits (composition, structure, organic matter, landscape position and micronutrients concentration).

The introduction of beneficial microorganisms through crop inoculation is now becoming popular among farmers, both in organic farming and unconventional agricultural practices. Nevertheless, the tuning of inoculants consistent with indigenous PGPM residents in the soil still faces a challenge in precision agriculture (Thilakarathna and Raizada, 2018).

Actually, the success of microbial inoculants' performance can be increased by precision farming methods, such as fertigation, which delivers the bioinoculants in the root area, thereby minimizing loss and interference with other liquid chemical inputs such as fertilizers, herbicides, pesticides and growth hormones, still used periodically (Bharathi et al., 2017). Furthermore, seed coating is widely used in agricultural practice, thanks to its beneficial impact on seed performance and plant productivity. Among coating agents, chemical pesticides, micronutrients, bio-stimulants and marker substances (tracing seeds within the crop supply chain) can be included. However, microbial (including bacteria and fungi) seed coating has been considered a cheap and efficient method for the delivery of inoculants. Recent research in precision agriculture has mainly focused on the optimization of such coating, setting up formulation of multi-effect microbial consortia designed for the already targeted annual crops (such as cereals, legumes, and some vegetables) and/or developed for other agricultural products.

However, it is becoming clear that, apart from soil and climatic features, which determine the choice of a specific plant variety by the farmer, the symbiotic rhizobia (as the other inoculants) genotypes matter. As reported above, GxG effects are present, determining the competitive abilities of the inoculant against the indigenous rhizobial populations. As for genetic screening and breeding programs for crops, there is a need for similar strategies for the symbiotic rhizobia. Several collections of rhizobia are present in the world, such as those at the USDA in the USA, the Centre for Rhizobium Studies (CRS) in Australia or at ARRIAM in the Russian Federation. Genome sequencing programs (e.g., GEBA at DOE-JGI, USA) and analysis of strain genotypes have been conducted. For instance, in *S. meliloti*, more than 280 high quality genome sequences are available on the GenBank database (October 2020). However, informa-

tion on the set of genes linked to the variation in symbiotic partnership performances (such as GxG) are still in their infancy. We need more comprehensive, holobiont-centered and systems biology-oriented studies to provide the basis for a future integration into plant breeding programs, concurrent symbiotic "breeding" programs, resulting in outperformance of the symbiotic partnership.

References

Adesemoye, A. O., Torbert, H. A., and Kloepper, J. W. 2009. Plant growth-promoting rhizobacteria allow reduced application rates of chemical fertilizers. *Microbial Ecology 58*(4):921–929. https://doi.org/10.1007/s00248-009-9531-y

Bakker, M. G., Manter, D. K., Sheflin, A. M., Weir, T. L., and Vivanco, J. M. 2012. Harnessing the rhizosphere microbiome through plant breeding and agricultural management. *Plant and Soil* 360(1–2):1–13. https://doi.org/10.1007/s11104-012-1361-x

Barnes, D., Heichel, G., Vance, C., and Ellis, W. 1984. A multiple-trait breeding program for improving the symbiosis for N_2 fixation between *Medicago sativa* L. and *Rhizobium meliloti*. *Plant and Soil* 82(3):303–314. https://doi.org/10.1007/BF02184269

Bellabarba, A., Fagorzi, C., diCenzo, G. C., Pini, F., Viti, C., and Checcucci, A. 2019. Deciphering the symbiotic plant microbiome: translating the most recent discoveries on rhizobia for the improvement of agricultural practices in metal-contaminated and high saline lands. *Agronomy* 9(9):529. https://doi.org/10.3390/agronomy9090529

Bharathi, J., Balachander, D., Kumar, K., and Narayanan, R. 2017. Evaluation of new microbial consortium through biofertigation for precision farming of bhendi (COBH 1). *International Journal of Medical Sciences and Pharmaceutical Research* 1(1):15–24.

Bliss, F. 1990. Utilization of genetic resources for crop improvement: The common bean. In: Brown, A. H. D., Clegg, M. T., Kahler, A. L., and Weir, B. S. (Eds.), Plant population genetics, breeding, and genetic resources, pp. 317–333, Sinauer, Sunderland, MA.

Bloemberg, G. V. and Lugtenberg, B. J. J. 2001. Molecular basis of plant growth promotion and biocontrol by rhizobacteria. *Current Opinion in Plant Biology* 4(4):343–350. https://doi.org/10.1016/S1369-5266(00)00183-7

Bose, J. and Schulte, R. D. 2014. Testing GxG interactions between coinfecting microbial parasite genotypes within hosts. *Frontiers in Genetics* 5:124. https://doi.org/10.3389/fgene.2014.00124

Bozsoki, Z., Gysel, K., Hansen, S. B., Lironi, D., Krönauer, C., Feng, F., de Jong, N., Vinther, M., Kamble, M., Thygesen, M. B., Engholm, E., Kofoed, C., Fort, S., Sullivan, J. T., Ronson, C. W., Jensen, K. J., Blaise, M., Oldroyd, G., Stougaard, J., ... Radutoiu, S. 2020. Ligand-recognizing motifs in plant LysM receptors are major determinants of specificity. *Science* 369(6504):663–670. https://doi.org/10.1126/science.abb3377

Brinker, P., Fontaine, M. C., Beukeboom, L. W., and Falcao Salles, J. 2019. Host, symbionts, and the microbiome: The missing tripartite interaction. *Trends in Microbiology* 27(6):480–488. https://doi.org/10.1016/j.tim.2019.02.002

Bronstein, J. L. 2015. Mutualism. Oxford University Press, USA. https://doi.org/10.1093/acprof:oso/9780199675654.001.0001

Burghardt, L. T. 2020. Evolving together, evolving apart: Measuring the fitness of rhizobial bacteria in and out of symbiosis with leguminous plants. *New Phytologist* 228(1):28–34. https://doi.org/10.1111/nph.16045

Burghardt, L. T., Epstein, B., Guhlin, J., Nelson, M. S., Taylor, M. R., Young, N. D., Sadowsky, M. J., and Tiffin, P. 2018. Select and resequence reveals relative fitness of bacteria in symbiotic and free-living environments. *Proceedings of the National Academy of Sciences USA* 115(10):2425–2430. https://doi.org/10.1073/pnas.1714246115

Burghardt, L. T., Trujillo, D. I., Epstein, B., Tiffin, P., and Young, N. D. 2020. A select and resequence approach reveals strain-specific effects of *Medicago* nodule-specific PLAT-domain genes. *Plant Physiology* 182(1):463–471. https://doi.org/10.1104/pp.19.00831

Carelli, M., Gnocchi, S., Fancelli, S., Mengoni, A., Paffetti, D., Scotti, C., and Bazzicalupo, M. 2000. Genetic diversity and dynamics of *Sinorhizobium meliloti* populations nodulating different alfalfa cultivars in Italian soils. *Applied and Environmental Microbiology* 66(11):4785–4789. https://doi.org/10.1128/AEM.66.11.4785-4789.2000

Checcucci, A., Azzarello, E., Bazzicalupo, M., Galardini, M., Lagomarsino, A., Mancuso, S., Marti, L., Marzano, M. C., Mocali, S., Squartini, A., Zanardo, M., and Mengoni, A. 2016. Mixed nodule infection in *Sinorhizobium meliloti–Medicago sativa* symbiosis suggest the presence of cheating behavior. *Frontiers in Plant Science* 7:835. https://doi.org/10.3389/fpls.2016.00835

Checcucci, A., DiCenzo, G. C., Bazzicalupo, M., and Mengoni, A. 2017. Trade, diplomacy, and warfare: The quest for elite rhizobia inoculant strains. *Frontiers in Microbiology* 8:2207. https://doi.org/10.3389/fmicb.2017.02207

de Vries, F. T., Griffiths, R. I., Knight, C. G., Nicolitch, O., and Williams, A. 2020. Harnessing rhizosphere microbiomes for drought-resilient crop production. *Science* 368(6488):270–274. https://doi.org/10.1126/science.aaz5192

diCenzo, G. C., Tesi, M., Pfau, T., Mengoni, A., and Fondi, M. 2020. Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. *Nature Communications* 11(1):2574. https://doi.org/10.1038/s41467-020-16484-2

diCenzo, G. C., Zamani, M., Checcucci, A., Fondi, M., Griffitts, J. S., Finan, T. M., and Mengoni, A. 2019. Multidisciplinary approaches for studying rhizobium–legume symbioses. *Canadian Journal of Microbiology* 65(1):1–33. https://doi.org/10.1139/cjm-2018-0377

Dubey, R. K., Tripathi, V., Prabha, R., Chaurasia, R., Singh, D. P., Rao, C. S., El-Keblawy, A., and Abhilash, P. C. 2020. Unravelling the soil microbiome: Perspectives for environmental sustainability. Springer. https://doi.org/10.1007/978-3-030-15516-2

Escudero-Martinez, C. and Bulgarelli, D. 2019. Tracing the evolutionary routes of plant–microbiota interactions. *Current Opinion in Microbiology* 49:34–40. https://doi.org/10.1016/j.mib.2019.09.013

Fageria, N. and Baligar, V. 2005. Enhancing nitrogen use efficiency in crop plants. *Advances in Agronomy* 88:97–185. https://doi.org/10.1016/S0065-2113(05)88004-6

Fagorzi, C., Bacci, G., Huang, R., Cangioli, L., Checcucci, A., Fini, M., Perrin, E., Natali, C., diCenzo, G. C., and Mengoni, A. 2021. Non-additive transcriptomic signatures of genotype x genotype interactions during the initiation of plant-rhizobium symbiosis. *mSystems* 6(1):e00974-20. https://doi.org/10.1128/mSystems.00974-20

Geurts, R. and Bisseling, T. 2002. *Rhizobium* nod factor perception and signalling. *The Plant Cell* 14(suppl 1):S239. https://doi.org/10.1105/tpc.002451

Gibson, K. E., Kobayashi, H., and Walker, G. C. 2008. Molecular determinants of a symbiotic chronic infection. *Annual Re-*

view of Genetics 42(1):413–441. https://doi.org/10.1146/annurev.genet.42.110807.091427

Griesmann, M., Chang, Y., Liu, X., Song, Y., Haberer, G., Crook, M. B., Billault-Penneteau, B., Lauressergues, D., Keller, J., Imanishi, L., Roswanjaya, Y. P., Kohlen, W., Pujic, P., Battenberg, K., Alloisio, N., Liang, Y., Hilhorst, H., Salgado, M. G., Hocher, V., … Cheng, S. 2018. Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. *Science* 361(6398):eaat1743. https://doi.org/10.1126/science.aat1743

Hacquard, S., Garrido-Oter, R., González, A., Spaepen, S., Ackermann, G., Lebeis, S., McHardy, A. C., Dangl, J. L., Knight, R., Ley, R., and Schulze-Lefert, P. 2015. Microbiota and host nutrition across Plant and Animal Kingdoms. *Cell Host and Microbe* 17(5):603–616. https://doi.org/10.1016/j.chom.2015.04.009

Heath, K. D. 2008. The coevolutionary genetics of plant-microbe interactions. *New Phytologist* 180(2):268–270. https://doi.org/10.1111/j.1469-8137.2008.02633.x

Igolkina, A. A., Bazykin, G. A., Chizhevskaya, E. P., Provorov, N. A., and Andronov, E. E. 2019. Matching population diversity of rhizobial *nod* A and legume *NFR5* genes in plant–microbe symbiosis. *Ecology and Evolution* 9(18):10377–10386. https://doi.org/10.1002/ece3.5556

Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E., and Walker, G. C. 2007. How rhizobial symbionts invade plants: The *Sinorhizobium–Medicago* model. *Nature Reviews Microbiology* 5(8):619–633. https://doi.org/10.1038/nrmicro1705

Joseph, A., Chandra, J., and Siddharthan, S. 2020. Genome analysis for precision agriculture using artificial intelligence: A survey. In Data Science and Security, pp. 221–226. Springer. https://doi.org/10.1007/978-981-15-5309-7_23

Kang, W., Jiang, Z., Chen, Y., Wu, F., Liu, C., Wang, H., Shi, S., and Zhang, X.-X. 2020. Plant transcriptome analysis reveals specific molecular interactions between alfalfa and its rhizobial symbionts below the species level. *BMC Plant Biology* 20(1):293. https://doi.org/10.1186/s12870-020-02503-3

Ke, J., Wang, B., and Yoshikuni, Y. 2020. Microbiome engineering: synthetic biology of plant-associated microbiomes in sustainable agriculture. *Trends in Biotechnology*. https://doi.org/10.1016/j.tibtech.2020.07.008

Kereszt, A., Mergaert, P., and Kondorosi, E. 2011. Bacteroid development in legume nodules: Evolution of mutual benefit or of sacrificial victims? *Molecular Plant-Microbe Interactions* 24(11):1300–1309. https://doi.org/10.1094/MPMI-06-11-0152

Kiers, E. T. and Denison, R. F. 2008. Sanctions, cooperation, and the stability of plant-rhizosphere mutualisms. *Annual Review of Ecology, Evolution and Systematics* 39(1):215–236. https://doi.org/10.1146/annurev.ecolsys.39.110707.173423

Koskella, B. and Bergelson, J. 2020. The study of host–microbiome (co)evolution across levels of selection. *Philosophical Transactions of the Royal Society B: Biological Sciences* 375(1808):20190604. https://doi.org/10.1098/rstb.2019.0604

Lanier, K. A., Petrov, A. S., and Williams, L. D. 2017. The central symbiosis of molecular biology: Molecules in mutualism. *Journal of Molecular Evolution* 85(1–2):8–13. https://doi.org/10.1007/s00239-017-9804-x

Levy, A., Conway, J. M., Dangl, J. L., and Woyke, T. 2018. Elucidating bacterial gene functions in the plant microbiome. *Cell Host and Microbe* 24(4):475–485. https://doi.org/10.1016/j.chom.2018.09.005

Liu, S., Ratet, P., and Magne, K. 2020. Nodule diversity, evolution, organogenesis and identity. *Advances in Botanical Research* 94:119–148. https://doi.org/10.1016/bs.abr.2019.09.009

Lucy, M., Reed, E., and R. Glick, B. 2004. Applications of free living plant growth-promoting rhizobacteria. *Antonie van Leeuwenhoek* 86:1–25. https://doi.org/10.1023/B:ANTO.0000024903.10757.6e

Malfanova, N. 2013. Endophytic bacteria with plant growth promoting and biocontrol abilities. Doctoral Thesis, Leiden University.

Martínez-Romero, E., Aguirre-Noyola, J. L., Taco-Taype, N., Martínez-Romero, J., and Zuñiga-Dávila, D. 2020. Plant microbiota modified by plant domestication. *Systematic and Applied Microbiology* 43(5):126106. https://doi.org/10.1016/j.syapm.2020.126106

Moran, N. A. and Sloan, D. B. 2015. The hologenome concept: helpful or hollow? *PLOS Biology* 13(12):e1002311. https://doi.org/10.1371/journal.pbio.1002311

Paffetti, D., Daguin, F., Fancelli, S., Gnocchi, S., Lippi, F., Scotti, C., and Bazzicalupo, M. 1998. Influence of plant genotype on the selection of nodulating *Sinorhizobium meliloti* strains by *Medicago sativa*. *Antonie van Leeuwenhoek* 73:3–8. https://doi.org/10.1023/A:1000591719287

Pawlowski, K. and Newton, W. E. 2008. Nitrogen-fixing actinorhizal symbioses. Springer. https://doi.org/10.1007/978-1-4020-3547-0

Pislariu, C. I., Sinharoy, S., Torres-Jerez, I., Nakashima, J., Blancaflor, E. B., and Udvardi, M. K. 2019. The nodule-specific PLAT domain protein NPD1 is required for nitrogen-fixing symbiosis. *Plant Physiology* 180(3):1480–1497. https://doi.org/10.1104/pp.18.01613

Provorov, N. A. and Tikhonovich, I. A. 2003. Genetic resources for improving nitrogen fixation in legume-rhizobia symbiosis. *Genetic Resources and Crop Evolution* 50(1):89–99. https://doi.org/10.1023/A:1022957429160

Pueppke, S. G. and Broughton, W. J. 1999. *Rhizobium* sp. strain NGR234 and *R. fredii* USDA257 share exceptionally broad, nested host ranges. *Molecular Plant-Microbe Interactions* 12(4):293–318. https://doi.org/10.1094/MPMI.1999.12.4.293

Qiu, Z., Egidi, E., Liu, H., Kaur, S., and Singh, B. K. 2019. New frontiers in agriculture productivity: Optimised microbial inoculants and in situ microbiome engineering. *Biotechnology Advances* 37(6):107371. https://doi.org/10.1016/j.biotechadv.2019.03.010

Remigi, P., Zhu, J., Young, J. P. W., and Masson-Boivin, C. 2016. Symbiosis within symbiosis: Evolving nitrogen-fixing legume symbionts. *Trends in Microbiology* 24(1):63–75. https://doi.org/10.1016/j.tim.2015.10.007

Rosenberg, E. and Zilber-Rosenberg, I. 2016. Microbes drive evolution of animals and plants: The hologenome concept. *MBio* 7(2):e01395-15. https://doi.org/10.1128/mBio.01395-15

Roy, S., Liu, W., Nandety, R. S., Crook, A., Mysore, K. S., Pislariu, C. I., Frugoli, J., Dickstein, R., and Udvardi, M. K. 2020. Celebrating 20 years of genetic discoveries in legume nodulation and symbiotic nitrogen fixation. *The Plant Cell* 32(1):15–41. https://doi.org/10.1105/tpc.19.00279

Sachs, J. L., Quides, K. W., and Wendlandt, C. E. 2018. Legumes versus rhizobia: A model for ongoing conflict in symbiosis. *New Phytologist* 219(4):1199–1206. https://doi.org/10.1111/nph.15222

Schlaeppi, K. and Bulgarelli, D. 2015. The plant microbiome at work. *Molecular Plant-Microbe Interactions* 28(3):212–217. https://doi.org/10.1094/MPMI-10-14-0334-FI

Shakoor, N., Northrup, D., Murray, S., and Mockler, T. C. 2019. Big data driven agriculture: Big data analytics in plant breeding, genomics, and the use of remote sensing technologies to advance crop productivity. *The*

Plant Phenome Journal 2(1):1–8. https://doi.org/10.2135/tppj2018.12.0009

Singh, B. K., Bardgett, R. D., Smith, P., and Reay, D. S. 2010. Microorganisms and climate change: Terrestrial feedbacks and mitigation options. *Nature Reviews Microbiology* 8(11):779–790. https://doi.org/10.1038/nrmicro2439

Soumare, A., Diedhiou, A. G., Thuita, M., Hafidi, M., Ouhdouch, Y., Gopalakrishnan, S., and Kouisni, L. 2020. Exploiting biological nitrogen fixation: A route towards a sustainable agriculture. *Plants* 9(8):1011. https://doi.org/10.3390/plants9081011

Sudheer, S., Bai, R. G., Usmani, Z., and Sharma, M. 2020. Insights on engineered microbes in sustainable agriculture: Biotechnological developments and future prospects. *Current Genomics* 21(5):321–333. https://doi.org/10.2174/1389202921999200603165934

Sulima, A. S., Zhukov, V. A., Afonin, A. A., Zhernakov, A. I., Tikhonovich, I. A., and Lutova, L. A. 2017. Selection signatures in the first exon of paralogous receptor kinase genes from the Sym2 region of the *Pisum sativum* L. genome. *Frontiers in Plant Science* 8:1957. https://doi.org/10.3389/fpls.2017.01957

Thilakarathna, M. S. and Raizada, M. N. 2018. Challenges in using precision agriculture to optimize symbiotic nitrogen fixation in legumes: Progress, limitations, and future improvements needed in diagnostic testing. *Agronomy* 8(5):78. https://doi.org/10.3390/agronomy8050078

Tirichine, L., de Billy, F., and Huguet, T. 2000. *Mtsym6*, a gene conditioning *Sinorhizobium* strain-specific nitrogen fixation in *Medicago truncatula*. *Plant Physiology* 123(3):845–852. https://doi.org/10.1104/pp.123.3.845

Turner, T. R., James, E. K., and Poole, P. S. 2013. The plant microbiome. *Genome Biology* 14(6):209. https://doi.org/10.1186/gb-2013-14-6-209

Wade, M. J. 2007. The co-evolutionary genetics of ecological communities. *Nature Reviews Genetics* 8(3):185–195. https://doi.org/10.1038/nrg2031

Wang, Q., Liu, J., Li, H., Yang, S., Körmöczi, P., Kereszt, A., and Zhu, H. 2018. Nodule-specific cysteine-rich peptides negatively regulate nitrogen-fixing symbiosis in a strain-specific manner in *Medicago truncatula*. *Molecular Plant-Microbe Interactions* 31(2):240–248. https://doi.org/10.1094/MPMI-08-17-0207-R

Werner, G. D. A., Strassmann, J. E., Ivens, A. B. F., Engelmoer, D. J. P., Verbruggen, E., Queller, D. C., Noe, R., Johnson, N. C., Hammerstein, P., and Kiers, E. T. 2014. Evolution of microbial markets. *Proceedings of the National Academy of Sciences USA* 111(4):1237–1244. https://doi.org/10.1073/pnas.1315980111

Westhoek, A., Field, E., Rehling, F., Mulley, G., Webb, I., Poole, P. S., and Turnbull, L. A. 2017. Policing the legume-*Rhizobium* symbiosis: A critical test of partner choice. *Scientific Reports* 7(1):1419. https://doi.org/10.1038/s41598-017-01634-2

Wilson, D. S. and Sober, E. 1989. Reviving the superorganism. *Journal of Theoretical Biology* 136(3):337–356. https://doi.org/10.1016/S0022-5193(89)80169-9

Yamaya-Ito, H., Shimoda, Y., Hakoyama, T., Sato, S., Kaneko, T., Hossain, M. S., Shibata, S., Kawaguchi, M., Hayashi, M., Kouchi, H., and Umehara, Y. 2018. Loss-of-function of ASPARTIC PEPTIDASE NODULE — INDUCED 1 (APN 1) in *Lotus japonicus* restricts efficient nitrogen-fixing symbiosis with specific *Mesorhizobium loti* strains. *The Plant Journal* 93(1):5–16. https://doi.org/10.1111/tpj.13759

Yang, X. and Fang, S. 2015. Practices, perceptions, and implications of fertilizer use in East-Central China. *Ambio* 44(7):647–652. https://doi.org/10.1007/s13280-015-0639-7

Zheng, M., Zhou, Z., Luo, Y., Zhao, P., and Mo, J. 2019. Global pattern and controls of biological nitrogen fixation under nutrient enrichment: A meta-analysis. *Global Change Biology* 25(9):3018–3030. https://doi.org/10.1111/gcb.14705

Zilber-Rosenberg, I. and Rosenberg, E. 2008. Role of microorganisms in the evolution of animals and plants: The hologenome theory of evolution. *FEMS Microbiology Reviews* 32(5):723–735. https://doi.org/10.1111/j.1574-6976.2008.00123.x

Symbiotic and Nonsymbiotic Members of the Genus *Ensifer* (syn. *Sinorhizobium*) Are Separated into Two Clades Based on Comparative Genomics and High-Throughput Phenotyping

Camilla Fagorzi[1], Alexandru Ilie[1], Francesca Decorosi[2], Lisa Cangioli[1], Carlo Viti[2], Alessio Mengoni [1,*], and George C. diCenzo [1,3,*]

[1]Department of Biology, University of Florence, Sesto Fiorentino, Italy

[2]Genexpress Laboratory, Department of Agriculture, Food, Environment and Forestry, University of Florence, Sesto Fiorentino, Italy

[3]Department of Biology, Queen's University, Kingston, Ontario, Canada

*Corresponding authors: E-mails: alessio.mengoni@unifi.it; george.dicenzo@queensu.ca.

Accepted: 12 October 2020

Data deposition: Genome sequences have been deposited at NCBI under the BioProject accession PRJNA622509.

Abstract

Rhizobium–legume symbioses serve as paradigmatic examples for the study of mutualism evolution. The genus *Ensifer* (syn. *Sinorhizobium*) contains diverse plant-associated bacteria, a subset of which can fix nitrogen in symbiosis with legumes. To gain insights into the evolution of symbiotic nitrogen fixation (SNF), and interkingdom mutualisms more generally, we performed extensive phenotypic, genomic, and phylogenetic analyses of the genus *Ensifer*. The data suggest that SNF likely emerged several times within the genus *Ensifer* through independent horizontal gene transfer events. Yet, the majority (105 of 106) of the *Ensifer* strains with the *nodABC* and *nifHDK* nodulation and nitrogen fixation genes were found within a single, monophyletic clade. Comparative genomics highlighted several differences between the "symbiotic" and "nonsymbiotic" clades, including divergences in their pangenome content. Additionally, strains of the symbiotic clade carried 325 fewer genes, on average, and appeared to have fewer rRNA operons than strains of the nonsymbiotic clade. Initial characterization of a subset of ten *Ensifer* strains identified several putative phenotypic differences between the clades. Tested strains of the nonsymbiotic clade could catabolize 25% more carbon sources, on average, than strains of the symbiotic clade, and they were better able to grow in LB medium and tolerate alkaline conditions. On the other hand, the tested strains of the symbiotic clade were better able to tolerate heat stress and acidic conditions. We suggest that these data support the division of the genus *Ensifer* into two main subgroups, as well as the hypothesis that pre-existing genetic features are required to facilitate the evolution of SNF in bacteria.

Key words: mutualism, evolutionary biology, phenomics, comparative genomics, rhizobia, Proteobacteria.

Significance

The bacterial genus *Ensifer* contains ecologically important N_2-fixing symbionts of leguminous plants, as well as nonsymbiotic species. However, the evolutionary dynamics of symbiotic nitrogen fixation within this genus are unclear, and it remains an open question of whether the gain of classical symbiotic N_2-fixation genes is sufficient to allow a bacterium to fix nitrogen. Our results suggest that the symbiotic species of the genus *Ensifer* predominately group separately from the nonsymbiotic species, but that symbiotic abilities were likely acquired multiple times within this group. This study provides new insight into the evolution of symbiotic N_2-fixation in a bacterial genus, while supporting the hypothesis that genetic features aside from the classical symbiotic N_2-fixation genes contribute to the evolution of symbiotic potential.

© The Author(s) 2020. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

Introduction

Symbioses are pervasive phenomena present in all Eukaryotic forms of life (López-García et al. 2017). These includes facultative symbiotic interactions, obligate symbioses, and the evolution of organelles (Douglas 2014), with symbiotic nitrogen fixation (SNF) being a paradigmatic example of the latter (Masson-Boivin and Sachs 2018). SNF (the conversion of N_2 to NH_3) is performed by a polyphyletic group of bacteria from the Alphaproteobacteria and Betaproteobacteria (whose nitrogen-fixing members are collectively called rhizobia) and members of the genus *Frankia* (Masson-Boivin et al. 2009; Wang and Young 2019) while intracellularly housed within specialized organs (nodules) of specific plants in the family *Fabaceae* and the genus *Parasponia*, as well as the actinorhizal plants (Werner et al. 2014; Griesmann et al. 2018; van Velzen et al. 2018). The advantages and evolutionary constraints to SNF have long been investigated in the conceptual framework of mutualistic interactions and the exchange of goods (see, for instance, Heath and Tiffin 2007; Werner et al. 2015; Sørensen et al. 2019), and quantitative estimations with metabolic reconstructions have also been performed (Pfau et al. 2018; diCenzo et al. 2020).

The establishment of a symbiotic nitrogen-fixing interaction requires that the bacterium encode several diverse molecular functions, including those related to signaling and metabolic exchange with the host plant, nitrogenase and nitrogenase-related functions, and escaping or resisting the plant immune system (Oldroyd et al. 2011; Haag et al. 2013; Poole et al. 2018). In general, the primary genes required for SNF (i.e., the *nod*, *nif*, and *fix* genes) are located within mobile genetic elements that include symbiotic islands and symbiotic (mega)-plasmids (Checcucci et al. 2019; Tian and Young 2019; Geddes et al. 2020), facilitating their spread through horizontal gene transfer (HGT) (Sullivan et al. 1995; Barcellos et al. 2007; Pérez Carrascal et al. 2016). Emphasizing the role of HGT in the evolution of rhizobia, rhizobia are dispersed across seven families of the Alphaproteobacteria and one family of the Betaproteobacteria, and most genera with rhizobia also contain nonrhizobia (Garrido-Oter et al. 2018; Wang 2019).

An interesting area of investigation is whether the evolution of mutualistic symbioses, such as SNF, depends on metabolic/genetic requirements ("facilitators," as in Gerhart and Kirschner [2007]) aside from the strict symbiotic genes (Long 2001; Sanjuán 2016; Zhao et al. 2018). In other words, 1) is the acquisition of symbiotic genes present in genomic islands or plasmids sufficient to become a symbiont? or, 2) are metabolic prerequirements or adaptation successive to HGT required? A comparative genomics study of 1,314 Rhizobiales genomes identified no functional difference between rhizobia and nonrhizobia based on Kyoto Encyclopedia of Genes and Genomes annotations (Garrido-Oter et al. 2018), suggestive of an absence of obvious facilitators. In contrast, experimental studies are generally consistent with an important role of nonsymbiotic genes in establishing or optimizing rhizobium–legume symbioses. Several studies have shown that effective symbionts are not produced following the transfer of symbiotic plasmids from rhizobia of the genera *Rhizobium* or *Ensifer* (syn. *Sinorhizobium*) to closely related nonrhizobia from the genera *Agrobacterium* or *Ensifer* (see, for instance, Hooykaas et al. 1982; Finan et al. 1986; Rogel et al. 2001; reviewed in diCenzo et al. [2019]). Similarly, the same symbiotic island is associated with vastly different symbiotic phenotypes depending on the *Mesorhizobium* genotype (Nandasena et al. 2007; Haskett et al. 2016). Further supporting the need for additional adaptations to support SNF, symbiosis plasmid transfer coupled to experimental evolution can lead to the gain of more advanced symbiotic phenotypes (Doin de Moura et al. 2020).

The genus *Ensifer* provides an ideal model to further explore the differentiation, or lack thereof, of symbiotic bacteria from nonsymbionts. This genus comprises rhizobia such as *Ensifer meliloti* and *Ensifer fredii*, as well as nonrhizobia like *Ensifer morelense* and *Ensifer adhaerens*, and many members have been extensively studied producing an abundant set of experimental and genomic data (for a recent review, see diCenzo et al. [2019]). The genus *Ensifer*, as currently defined, resulted from the combination of the genera *Sinorhizobium* and *Ensifer* based on similarities in the 16S rRNA and *recA* sequences of the type strains and the priority of the name *Ensifer* (Willems et al. 2003; Young 2003). Multilocus sequence analysis supported the amalgamation of these genera (Martens et al. 2007), although it was subsequently noted that *E. adhaerens* (the type strain) is an outgroup of this taxon based on whole genome phylogenomics (Ormeño-Orrillo et al. 2015). A more recent taxonomy approach based on genome phylogeny suggests that the genus *Ensifer* should again be split, with the initial type strains of *Ensifer* and *Sinorhizobium* belonging to separate genera (Parks et al. 2018).

In this article, we report an extensive comparative genomic analysis, and initial phenotypic characterization, of legume symbionts and nonsymbionts of the genus *Ensifer*. We identified that SNF likely evolved multiple times through independent HGT events; even so, most symbionts were found in a single clade, consistent with a requirement for pre-existing genetic features to facilitate the evolution of SNF. Moreover, the symbiotic and nonsymbiotic clades differed in their pangenome composition, and tests with a subset of strains suggested they also differed in their substrate utilization and resistance phenotypes as measured by the Phenotype MicroArray platform. We suggest that the data

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

support the division of the genus *Ensifer* into two subgroups, corresponding to the genera *Ensifer* and *Sinorhizobium* of the Genome Taxonomy Database (Parks et al. 2018).

Materials and Methods

Genome Sequencing, Assembly, and Annotation

Prior to short-read sequencing, all strains were grown to stationary phase at 30 °C in TY medium (5 g l^{-1} tryptone, 3 g l^{-1} yeast extract, and 0.4 g l^{-1} $CaCl_2$). Total genomic DNA was isolated using a standard cetyltrimethylammonium bromide method (Perrin et al. 2015). Short-read sequencing was performed at IGATech (Udine, Italy) using an Illumina HiSeq2500 instrument with 125-bp paired-end reads. Two independent sequencing runs were performed for *E. morelense* Lc04 and *Ensifer psoraleae* CCBAU 65732, whereas *E. morelense* Lc18 and *Ensifer sesbaniae* CCBAU 65729 were sequenced once. For long-read sequencing, *E. sesbaniae* was grown to midexponential phase at 30 °C in MM9 minimal medium (MOPS buffer [40 mM MOPS, 20 mM KOH], 19.2 mM NH_4Cl, 8.76 mM NaCl, 2 mM KH_2PO_4, 1 mM $MgSO_4$, 0.25 mM $CaCl_2$, 1 μg ml^{-1} biotin, 42 nM $CoCl_2$, 38 μM $FeCl_3$, 10 μM thiamine-HCl, and 10 mM sucrose). Total genomic DNA was isolated as described elsewhere (Cowie et al. 2006). Long-read sequencing was performed in-house with a Pacific Biosciences Sequel instrument.

Reads were assembled into scaffolds using SPAdes 3.9.0 (Bankevich et al. 2012; Vasilinetc et al. 2015); in the case of *E. sesbaniae*, long reads were corrected and trimmed using Canu 1.7.1 (Koren et al. 2017) prior to assembly. Scaffolds returned by SPAdes were parsed to remove those with <20× coverage or with a length <200 nucleotides. Using FastANI (Jain et al. 2018), one-way average nucleotide identity (ANI) values of each assembly were calculated against 887 alphaproteobacterial genomes available through the National Center for Biotechnological Information (NCBI) with an assembly level of complete or chromosome. Based on the FastANI output, each draft genome assembly was further scaffolded using MeDuSa (Bosi et al. 2015) and the reference genomes listed in supplementary table S1, Supplementary Material online. For most assemblies, scaffolds under 1 kb in length were discarded. The exception was for *S. sesbaniae*, for which case scaffolds <10 kb were discarded. Genome assemblies were annotated using Prokka 1.12-beta (Seemann 2014), annotating coding regions with Prodigal (Hyatt et al. 2010), tRNA with Aragon (Laslett and Canbäck 2004), rRNA with Barrnap (https://github.com/tseemann/barrnap; last accessed October 18, 2020), and ncRNA with Infernal (Kolbe and Eddy 2011) and Rfam (Kalvari et al. 2018).

Species Phylogenetic Analyses

All *Ensifer* (and *Sinorhizobium*) genomes were downloaded from the NCBI Genome Database regardless of assembly level. Strains that either 1) lacked a RefSeq assembly, 2) had genome sizes <1 Gb, or 3) appeared to not belong to the *Ensifer* clade based on preliminary phylogenetic analyses were discarded, leaving a final set of 157 strains (supplementary data set S1, Supplementary Material online). Eight complete *Rhizobium* genomes (supplementary data set S2, Supplementary Material online) were downloaded to serve as an outgroup. Genomes were reannotated with prokka to ensure consistent annotation. All genomes were downloaded on November 12, 2018, and associated metadata are available as supplementary data sets S1 and S2, Supplementary Material online.

To construct an unrooted, core gene phylogeny, the pangenome of the 157 *Ensifer* strains was calculated using Roary 3.11.3 (Page et al. 2015) with a percent identify threshold of 70%. As part of the running of Roary, the nucleotide sequences of the 1,049 core genes (identified as those found in at least 99% of the genomes; supplementary data set S3, Supplementary Material online) were individually aligned with PRANK (Löytynoja 2014) and the alignments concatenated. The concatenated alignment was trimmed using TRIMAL 1.2rev59 (Capella-Gutiérrez et al. 2009) with the automated1 option, and used to construct a maximum likelihood phylogeny (the bootstrap best tree following 100 bootstrap replicates, as determined by the extended majority-rule consensus tree criterion) using RAxML 8.2.9 (Stamatakis 2014) with the GTRCAT model as recommended (https://cme.h-its.org/exelixis/resource/download/NewManual.pdf; last accessed October 18, 2020). All phylogenies prepared in this study were visualized with the online iTOL webserver (Letunic and Bork 2016).

To construct a rooted phylogeny, the AMPHORA2 pipeline (Wu and Scott 2012) was used to identify 31 highly conserved bacterial proteins in each *Ensifer* and *Rhizobium* proteome, based on HMMER 3.1b2 (Eddy 2009) and the 31 hidden Markov models (HMMs) that come with AMPHORA2. Customs Perl scripts were then used to remove proteins that were either found in <95% of genomes or were found in multicopy in at least one genome, leaving a set of 30 proteins (Frr, InfC, NusA, Pgk, PyrG, RplA, RplB, RplC, RplD, RplE, RplF, RplK, RplL, RplM, RplN, RplP, RplS, RplT, RpmA, RpoB, RpsB, RpsC, RpsE, RpsI, RpsJ, RpsK, RpsM, RpsS, SmpB, Tsf). Orthologous groups were aligned using MAFFT 7.310 (Katoh and Standley 2013) with the localpair option, following which the alignments were trimmed using TRIMAL 1.2rev59 with the automated1 option. Alignments were concatenated and used to construct a maximum likelihood phylogeny (the bootstrap best tree following 304 bootstrap replicates, as determined by the extended majority-rule consensus tree criterion) using RAxML with the PROTGAMMAJTTDCMUT model. This model was chosen as a preliminary run using RAxML with the automatic model selection indicated that the best scoring tree was obtained with the selected model.

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

ANI and AAI Calculations

Pairwise ANI values were calculated for all *Ensifer* strains using FastANI (Jain et al. 2018) with default parameters; a value of 78% was used in cases where no value was returned by FastANI. Pairwise average amino acid identity (AAI) values were calculated with the compareM workflow (https://github.com/dparks1134/CompareM; last accessed October 18, 2020). Results were visualized and clustered using the heatmap.2 function of the gplots package in R (Warnes et al. 2016), with average linkage and Pearson correlation distances.

Pangenome Calculation

All proteins of the reannotated *Ensifer* strains were clustered into orthologous groups using CD-HIT 4.6 (Li and Godzik 2006) with a percent identity threshold of 70% and an alignment length of 80% of the longer protein. The output was used to determine core and accessory genomes using a prevalence threshold of 90% as many of the genomes were draft genomes. Gene accumulation curves were produced using the specaccum function of the vegan package of R (Oksanen et al. 2018), with the random method and 500 permutations. Principal component analysis (PCA) was performed with the prcomp function of R, and was visualized with the autoplot function the ggplot2 package (Wickham 2016).

Identification and Phylogenetic Analysis of Common Nod, Nif, and Rep Proteins

The proteomes were collected for the 157 *Ensifer* strains, as well as all strains from the genera *Rhizobium*, *Neorhizobium*, *Agrobacterium*, *Mesorhizobium*, and *Ochrobactrum* with an assembly status of Complete or Chromosome (supplementary data set S4, Supplementary Material online). Additionally, the seed alignments for the HMMs of the nodulation proteins NodA (TIGR04245), NodB (TIGR04243), and NodC (TIGR04242), the nitrogenase proteins NifH (TIGR01287), NifD (TIGR01282), NifK (TIGR01286), and the replicon partitioning proteins RepA (TIGR03453), and RepB (TIGR03454) were downloaded from TIGRFAM (Haft et al. 2012). Seed alignments were converted into HMMs with the HMMBUILD function of HMMER 3.1b2 (Eddy 2009). Each HMM was searched against the complete set of proteins from all 157 reannotated *Sinorhizobium* and *Ensifer* strains using the HMMSEARCH function of HMMER. The amino acid sequences for each hit (regardless of e-value) were collected. Each set of sequences was searched against a HMM database containing all 21,200 HMMs from the Pfam (Finn et al. 2016) and TIGRFAM databases using the HMMSCAN function of HMMER, and the top scoring HMM hit for each query protein was identified. Proteins were annotated as NodA, NodB, NodC, NifH, NifD, NifK, RepA, or RepB according to supplementary table S2, Supplementary Material online.

The *nodA*, *nodB*, and *nodC* genes are generally found as an operon. Thus, the NodA, NodB, and NodC proteins were putatively associated to operons based on identifying proteins that are encoded by adjacent genes in their respective genomes; orphan proteins not encoded by adjacent genes were discarded as the subsequent phylogenetic analysis was based on concatenated NodA, NodB, and NodC alignments. Each set of orthologs were aligned using MAFFT with the localpair option, and alignments trimmed using TRIMAL and the automated1 algorithm. Alignments were concatenated so as to combine alignments for proteins encoded by adjacent genes, producing a NodABC alignment. The same procedure was followed to produce NifHDK and RepAB alignments. Maximum likelihood phylogenies were built on the basis of each combined alignment using RAxML with the PROTGAMMAJTT (NodABC, NifHDK) or the PROTGAMMALG (RepAB) models. These models were chosen as preliminary runs using RAxML with the automatic model selection indicated that the best scoring trees were obtained with the selected models. The final phylogenies are the bootstrap best trees following 352 bootstrap replicates, as determined by the extended majority-rule consensus tree criterion.

Plant Assays

Phaseolus vulgaris (var. TopCrop, Mangani Sementi, Italy) seeds were surface sterilized in 2.5% $HgCl_2$ solution for 2 min and washed five times with sterile water. Seeds were germinated in the dark at 23 °C, following which seedlings were placed in sterile polypropylene jars containing vermiculite:perlite (1:1) and nitrogen-free Fåhraeus medium, and grown at 23 °C with a 12-h photoperiod (100 $\mu E\, m^{-2}\, s^{-1}$). One-week-old plantlets were inoculated with 100 µl of the appropriate rhizobium strain (suspended in 0.9% NaCl at an OD_{600} of 1); five plants were inoculated per strain and then grown for 4 weeks at 23 °C with a 12-h photoperiod (100 $\mu E\, m^{-2}\, s^{-1}$). Plant growth assays were repeated three independent times. Nodules were collected and surface sterilized as described elsewhere (Checcucci et al. 2016), crushed in sterile 0.9% NaCl solution, and serial dilutions were plated on TY agar plates and incubated at 30 °C for 2 days. PCR amplification of the 16S rRNA gene was performed using crude lysates from single colonies recovered from root nodules, as in Barzanti et al. (2007). Sequencing of the PCR amplified 16S rRNA gene was performed from both the 27f and the 1495r primers using BrilliantDye Terminator Cycle Sequencing chemistry (Nimagen, Nijmegen, The Netherlands) on a 3730xl DNA Analyzer (ThermoFisher Scientific, Waltham, MA).

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

Phenotype MicroArray

Phenotype MicroArray experiments using Biolog plates PM1 and PM2A (carbon sources), PM9 (osmolytes), and PM10 (pH) were performed as described previously (Biondi et al. 2009). Data were collected over 96 h with an OmniLog instrument. Data analysis was performed with DuctApe (Galardini et al. 2014). Activity index (AV) values were calculated following subtraction of the blank well from the experimental wells. Growth with each compound was evaluated with AV values from 0 (no growth) to 9 (maximal growth), following an elbow test calculation. Phenotype MicroArray experiments were performed once as results for these experiments are highly repeatable (Johnson et al. 2008; Bochner et al. 2010; Dunkley et al. 2019).

Biofilm Assays

Overnight cultures of strains grown in TY and LB (10 g l^{-1} tryptone, 5 g l^{-1} yeast extract, 5 g l^{-1} NaCl) media were diluted to an OD_{600} of 0.02 in fresh media, and six replicates of 100 µl aliquots were transferred to a 96-well microplate. Plates were incubated at 30 °C for 24 h, after which the OD_{600} was measured with a Tecan Infinite 200 PRO (Switzerland). Each well was then stained with 30 µl of a filtered 0.1% (w/v) crystal violet solution for 10 min, and then the medium containing the planktonic cells was gently removed from the wells. Next, the wells were rinsed three times with 200 µl of phosphate-buffered saline (PBS; 0.1 M, pH 7.4) and allowed to dry for 15 min. About 100 µl of 95% (v/v) ethanol was added to each well and then incubated for 15 min at room temperature. The OD_{540} of each well was measured (Rinaudi and González 2009), and biofilm production reported as the ratio of the OD_{540}/OD_{600} ratio. Biofilm assays consisted of six replicates, and were performed two independent times.

Growth Curves

Overnight cultures of each strain were grown in the same medium to be used for the growth curve. For minimal media, either 0.2% (w/v) of glucose or succinate was added as the carbon source. Cultures were diluted to an OD_{600} of 0.05 in the same media, and triplicate 150-µl aliquots were added to a 96-well microplate. Microplates were incubated without shaking at 30 °C or 37 °C in a Tecan Infinite 200 PRO, with OD_{600} readings taken every hour for 48 h. Growth rates were evaluated over 2-h windows during the exponential growth phase. All growth curves were performed in triplicate and repeated two independent times.

To evaluate bacterial growth when provided root exudates as a nitrogen source, root exudates were produced from *Medicago sativa* cv. Maraviglia as described elsewhere (Checcucci et al. 2017). Single bacterial colonies from TY plates were resuspended in a 0.9% NaCl solution to an OD_{600} of 0.5. Then, each well of a 96-well microplate was inoculated with 5 µl of culture, 75 µl of nitrogen-free M9 with 0.2% (w/v) succinate as a carbon source, and 20 µl of root exudate as a nitrogen source as done previously (Checcucci et al. 2017). Triplicates were performed for each strain. Microplates were incubated without shaking at 30 °C in a Tecan Infinite 200 PRO, with OD_{600} readings taken every hour for 48 h. Growth rates were determined as described above.

Results

Genome Sequencing of Four *Rhizobiaceae* Strains

Draft genomes of *E. morelense* Lc04, *E. morelense* Lc18, *E. sesbaniae* CCBAU 65729, and *E. psoraleae* CCBAU 65732 (Wang et al. 1999, 2013) were generated to increase the species diversity available for our analyses. Summary statistics of the assemblies are provided in supplementary table S3, Supplementary Material online. The genome sequences confirmed the presence of nodulation and nitrogen-fixing genes in *E. morelense* Lc18, *E. sesbaniae* CCBAU 65729, and *E. psoraleae* CCBAU 65732, whereas these genes appeared absent in the *E. morelense* Lc04 assembly. Strains Lc04, CCBAU 65729, and CCBAU 65732 were confirmed to belong to the genus *Ensifer*, as one-way ANI comparisons revealed that the most similar alpha-proteobacterial genomes were from the genus *Ensifer*. However, the genome of strain Lc18 was most similar to genomes from the genera *Rhizobium* and *Agrobacterium*, consistent with an earlier 16S rRNA gene restriction fragment length polymorphism analysis (34). Thus, we propose renaming *E. morelense* Lc18 to *Rhizobium* sp. Lc18. As this strain does not belong to the genus *Ensifer*, it was excluded from further analyses.

Symbiotic and Nonsymbiotic *Ensifer* Strains Segregate Phylogenetically

An unrooted, core gene phylogeny of 157 *Ensifer* strains was prepared to evaluate the phylogenetic relationships between the symbiotic and nonsymbiotic strains (fig. 1). A rooted phylogeny based on a multilocus sequence analysis was also prepared (supplementary fig. S1, Supplementary Material online). Each of the 157 strains was annotated as symbiotic or nonsymbiotic based on the presence of the common *nodABC* nodulation genes and the *nifHDK* nitrogenase genes. Consistent with previous work (Garrido-Oter et al. 2018), both phylogenies revealed a clear division of the symbiotic and nonsymbiotic strains into two well-defined clades. Nevertheless, a few exceptions were noted. *Ensifer sesbaniae* was found within the nonsymbiotic clade; however, *E. sesbaniae* was reported to be a symbiont of legumes such as *P. vulgaris* (Wang et al. 2013), and the ability of *E. sesbaniae* to

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

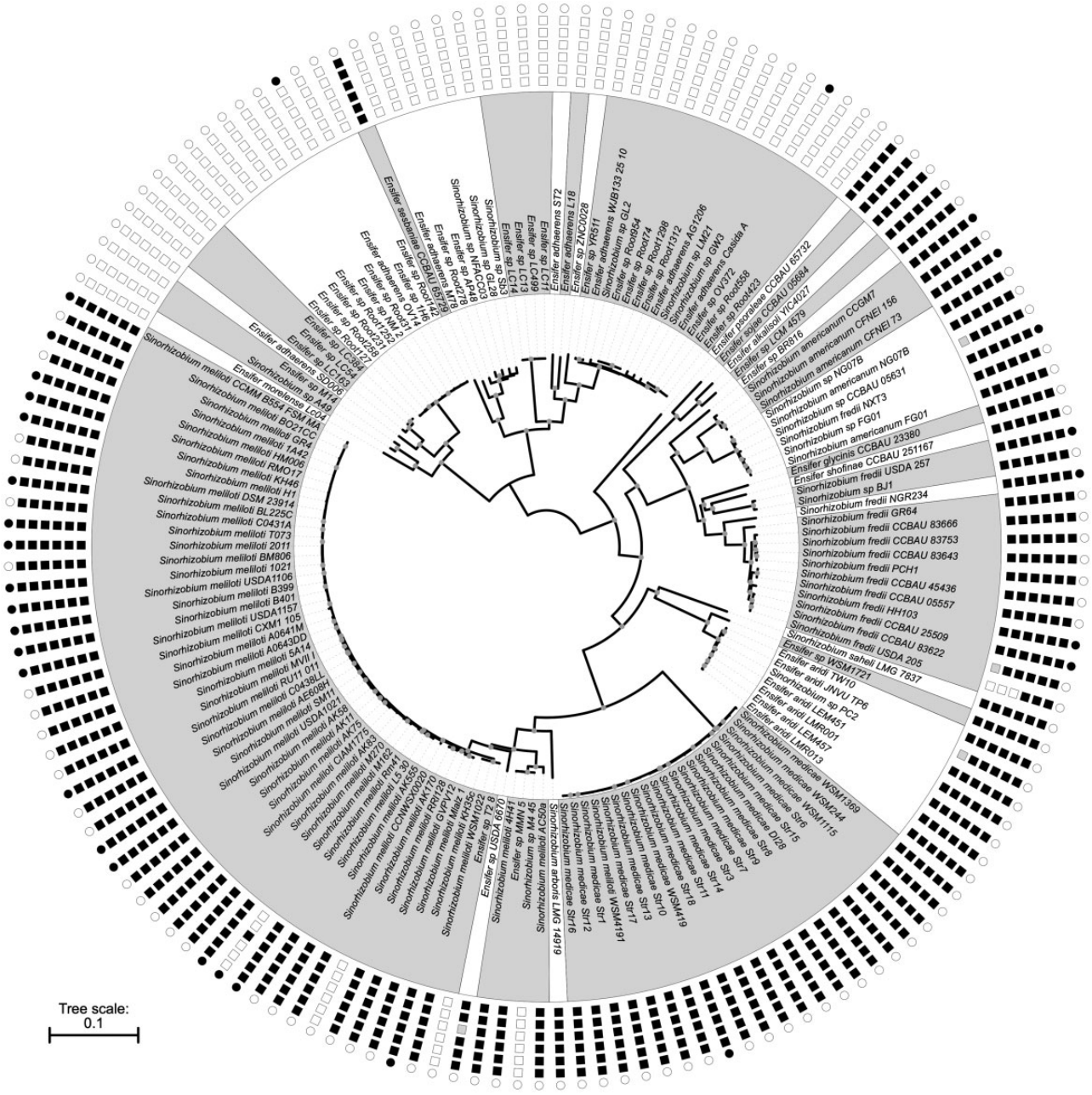


Fig. 1.—Unrooted phylogeny of the genus *Ensifer*. A maximum likelihood phylogeny of 157 strains was prepared from a concatenated alignment of 1,049 core genes. Nodes with a bootstrap value of 100 are indicated with the gray dots. The scale represents the mean number of nucleotide substitutions per site. The white and gray shading is used to group strains into genospecies on the basis of ANI and AAI results (supplementary figs. S3 and S4, Supplementary Material online), using ANI and AAI genospecies threshold of 95% and 96%, respectively. From outside to inside, rings represent the genome assembly level (black, finished; white, draft), and the presence (black) or absence (white) of NodA, NodB, NodC, NifH, NifD, and NifK. Gray boxes indicate the presence of a truncated version of the corresponding gene (as a result of incomplete genome assembly) detected through inspection of the RefSeq annotations. Strains are named as recorded in NCBI at the time of collection.

nodulate *P. vulgaris* was confirmed in this study (supplementary fig. S2, Supplementary Material online). Similarly, at least one of the six symbiotic proteins were not detected in five strains of the symbiotic group, although we cannot rule out that these are false negatives due to incomplete genome assemblies or genome assembly errors. ANI (genospecies threshold: 95%) and AAI (genospecies threshold: 96%) calculations suggested

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

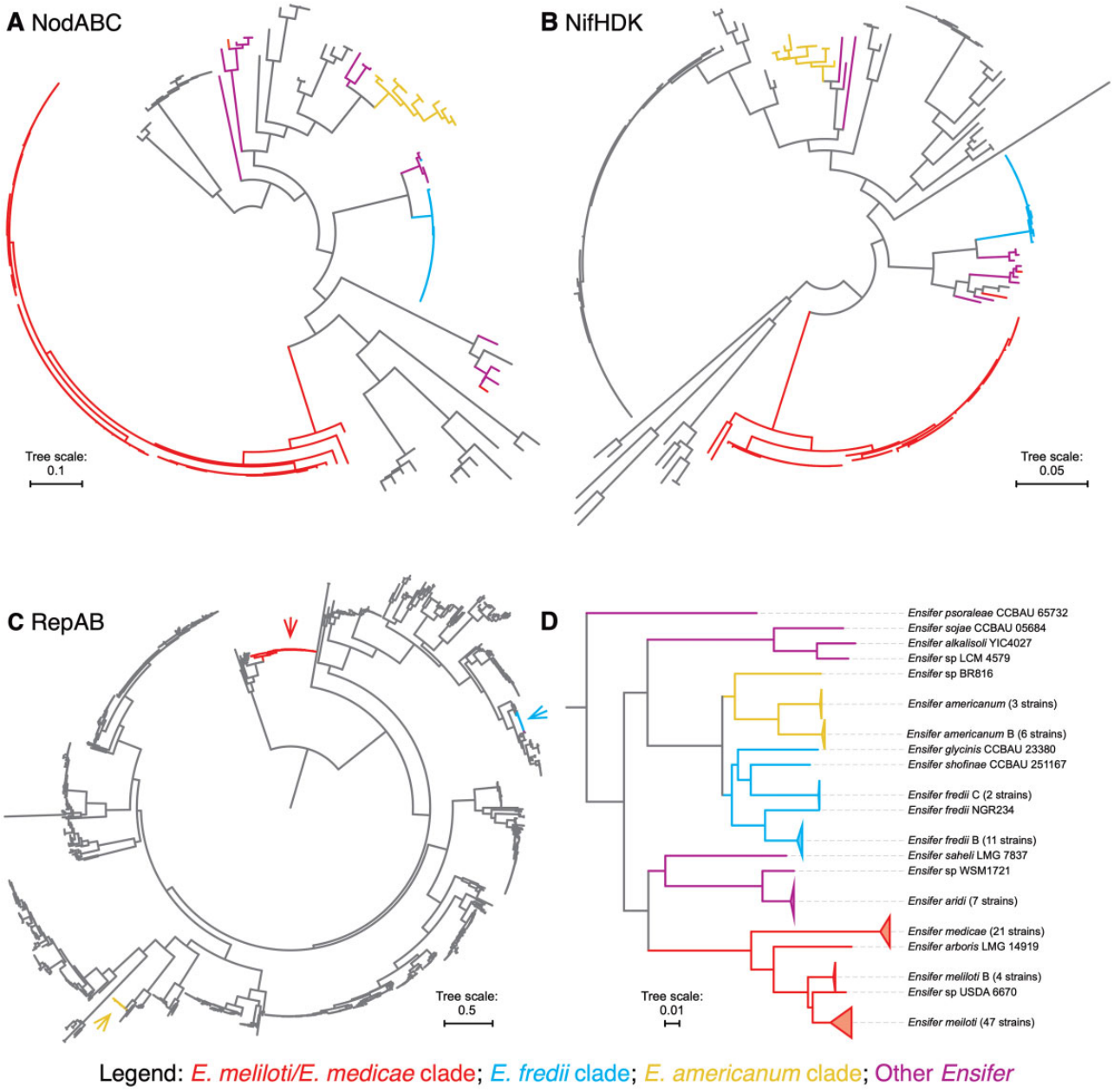


Fig. 2.—Evolution of SNF within the genus *Ensifer*. Maximum likelihood phylogenies of concatenated alignments of (*A*) NodABC nodulation proteins, (*B*) NifHDK nitrogenase proteins, and (*C*) RepAB replicon partitioning proteins of the order Rhizobiales. Branches corresponding to proteins from the genus *Ensifer* are indicated with color. (*D*) A subtree of the core gene species phylogeny of figure 1. Colors denote taxa whose symbiotic proteins are predicted to have been vertically acquired from a common ancestor. The scale bars represent the mean number of amino acid (*A*–*C*) or nucleotide (*D*) substitutions per site.

the presence of 12 and 20 genospecies within the nonsymbiotic and symbiotic groups, respectively (fig. 1 and supplementary figs. S3 and S4, Supplementary Material online), confirming that the nonsymbiotic clade is not an artifact of low species diversity. Thus, we conclude that the genus *Ensifer* consists of two well-defined clusters, each consisting predominately of either symbiotic or nonsymbiotic strains.

SNF Likely Arose Multiple Times within the Genus *Ensifer*

A possible explanation for the phylogenetic segregation of SNF within the genus *Ensifer* was that the symbiotic genes were gained once through a single HGT event. To test this hypothesis, the phylogenetic relationships of the NodABC and NifHDK proteins of the order Rhizobiales were examined (fig. 2*A* and *B*). SNF genes are situated on megaplasmids in

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

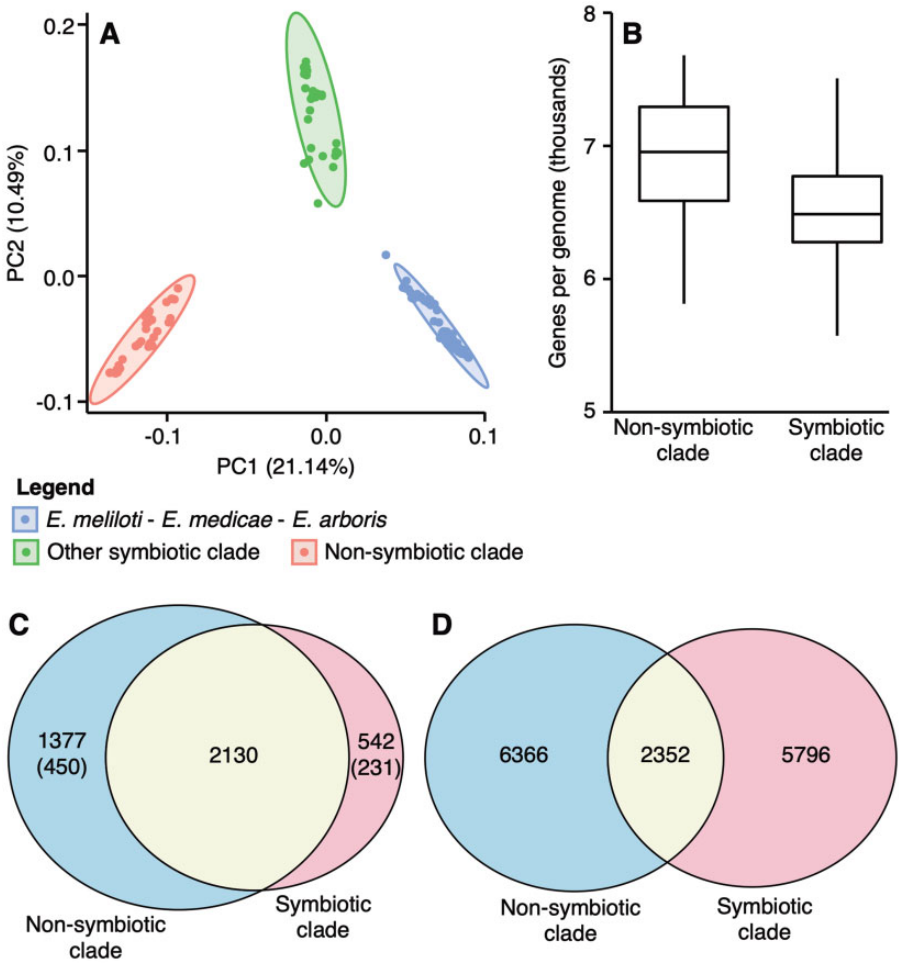


Fig. 3.—Global genome properties of the genus *Ensifer*. (*A*) A PCA plot based on the presence and absence of all orthologous protein groups in each of the 157 *Ensifer* strains. (*B*) Box-and-whisker plots displaying the number of genes per genome in the symbiotic and nonsymbiotic *Ensifer* clades. (*C*) A Venn diagram displaying the overlap in the core genomes of the symbiotic and nonsymbiotic *Ensifer* clades. (*D*) A Venn diagram displaying the overlap in the accessory genomes of the symbiotic and nonsymbiotic *Ensifer* clades.

the genus *Ensifer*; thus, a phylogeny of RepAB partitioning proteins of the order Rhizobiales was prepared as a proxy of the evolutionary relationships among the symbiotic megaplasmids (fig. 2*C*). We predicted that the NodABC, NifHDK, and RepAB proteins of the genus *Ensifer* would form a single, monophyletic clade in each phylogeny, if the above hypothesis were true. This was not observed. Instead, all three phylogenies were inconsistent with a single origin of SNF within the genus *Ensifer* as the *Ensifer* strains were predominantly split into three clades: 1) *E. meliloti* and *Ensifer medicae*, 2) *E. fredii* and related strains, and 3) *Ensifer americanum* and related strains. As the same clades are observed in the species tree (fig. 2*D*), this observation suggests SNF was independently acquired through HGT in each clade. The relationships between the SNF genes of the remaining *Ensifer* species (e.g., *Ensifer aridi* and *E. psoraleae*) was not clear; however, the most parsimonious solution was that there were three additional acquisitions of SNF (fig. 2*D*). Overall, the phylogenetic analyses of the NodABC, NifHDK, and RepAB proteins support the hypothesis that there were multiple, independent acquisition of symbiosis genes (hence SNF) by lineages within the genus *Ensifer*; however, the gain (and/or maintenance) of symbiosis genes preferentially occurred within one monophyletic group of species.

The Genomic Features of the Symbiotic and Nonsymbiotic Clades Differ

The pangenome of the 157 *Ensifer* strains was calculated to evaluate if there were global genomic differences between the symbiotic and nonsymbiotic clades. Both clades had open pangenomes (supplementary fig. S5, Supplementary Material online). A PCA based on gene presence/absence revealed a clear separation of the two clades (fig. 3*A*), suggesting a

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

Table 1

Ensifer Strains Phenotypically Characterized in This Study

Strain	Original Source	SNF[a]	*Ensifer* Clade[b]	Reference
Ensifer adhaerens Casida A	Isolated from a Pennsylvania soil sample	No	Nonsymbiotic	Casida (1982)
Ensifer adhaerens OV14	Isolated from the rhizosphere of *Brassica napus*	No	Nonsymbiotic	Wendt et al. (2012)
Ensifer sp. M14	Isolated from arsenic-rich sediments of a gold mine	No	Nonsymbiotic	Drewniak et al. (2008)
Ensifer morelense Lc04	Isolated from root nodules of *Leucaena leucocephala*	No	Nonsymbiotic	Wang et al. (1999)
Ensifer sesbaniae CCBAU 65729	Isolated from root nodules of *Sesbania cannabina*	Yes	Nonsymbiotic	Wang et al. (2013)
Ensifer fredii NGR234	Isolated from root nodules of *Lablab purpureus*	Yes	Symbiotic	Trinick (1980)
Ensifer sojae CCBAU 05684	Isolated from root nodules of *Glycine max* grown in saline-alkaline soils	Yes	Symbiotic	Li et al. (2011)
Ensifer americanum CFNEI 156	Isolated from root nodules of *Acacia acatlensis*	Yes	Symbiotic	Toledo et al. (2003)
Ensifer psoraleae CCBAU 65732	Isolated from root nodules of *Psoralea corylifolia*	Yes	Symbiotic	Wang et al. (2013)
Ensifer medicae WSM419	Isolated from root nodules of *Medicago murex*	Yes	Symbiotic	Howieson and Ewing (1986)

[a]This column indicates if the strain can (Yes) or cannot (No) form nitrogen-fixing nodules on legumes.

[b]This column indicates if the strain belongs to the symbiotic or nonsymbiotic clade of the genus *Ensifer* as defined in this study.

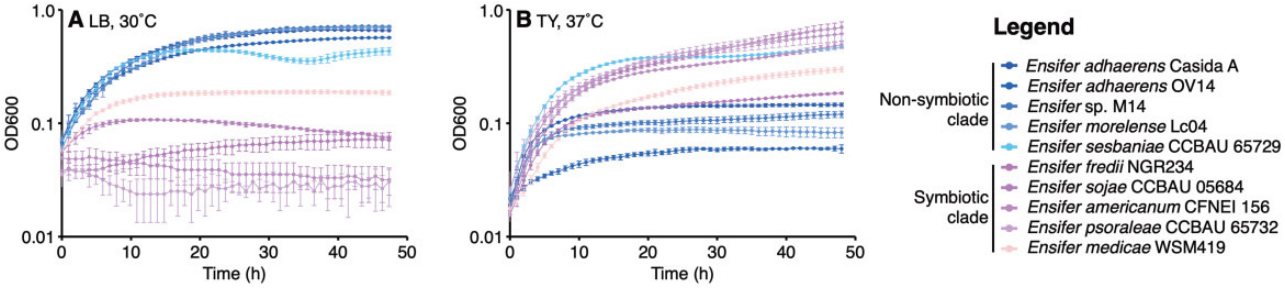


Fig. 4.—Growth properties of phylogenetically diverse *Ensifer* strains. *Ensifer* strains were grown in microplates without shaking. Data points represent the average of triplicate samples, whereas the error bars indicate the SD. Shades of pink are used for strains of the symbiotic clade, whereas shades of blue are used for strains of the nonsymbiotic clade. (*A*) Growth in LB medium at 30 °C. (*B*) Growth in TY medium during heat stress (37 °C).

divergence of the pangenomes of these clades. The symbiotic clade was subdivided into two groups along the second component of the PCA (fig. 3*A*), which may suggest further levels of genomic separation. About 2,130 genes were found in the core genomes of both clades, whereas 20% (542 genes) and 40% (1,377 genes) of the core genomes of the symbiotic and nonsymbiotic clades, respectively, were absent from the core genome of the other clade; of these, about a third were completely absent from the other clade's pangenome (fig. 3*C*). Of the 14,514 accessory genes (defined as genes found in at least 10% of at least one clade, excluding the 2,130 *Ensifer* core genes), only 2,352 (16%) were found in the pangenomes of both the symbiotic and the nonsymbiotic clades. Moreover, a statistically significant difference (Wilcoxon rank-sum test, $P < 0.0001$) in the genome sizes of the two clades was observed (fig. 3*B*); strains of the nonsymbiotic clade carried 325 more genes, on average, than strains of the symbiotic clade (median difference of 470). Finally, based on the limited number of strains with finished genomes, strains of the symbiotic clade appear to generally have three copies of the rRNA operon whereas strains of the nonsymbiotic clade appear to have a norm of five copies of their rRNA operon. Together, these multiple lines of data are consistent with there being a broad genomic divergence of the symbiotic and nonsymbiotic clades of the genus *Ensifer*.

Phenotypic Features of the Symbiotic and Nonsymbiotic Clades Differ

A subset of ten strains (table 1), five each from the symbiotic and nonsymbiotic clades, were subjected to a panel of assays to investigate how phenotypes vary across the genus *Ensifer*. These ten strains were chosen so as to provide broad phylogenetic coverage of the genus, while excluding strains for which extensive phenotypic characterizations have been previously published. No statistically significant differences were observed in the ability of members of the two clades to form biofilms (supplementary fig. S6 and table S4, Supplementary Material online). However, the tested strains clearly differed in their ability to grow in LB media; whereas the tested strains of the nonsymbiotic clade displayed robust growth in LB, the tested strains of the symbiotic clade largely failed to grow (fig. 4*A*). Tested strains of the nonsymbiotic clade also displayed a slightly faster specific growth rate, on average, than the tested strains of the symbiotic clade in TY media (nonsymbiotic clade: $0.54 \pm 0.03\,h^{-1}$; symbiotic clade:

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

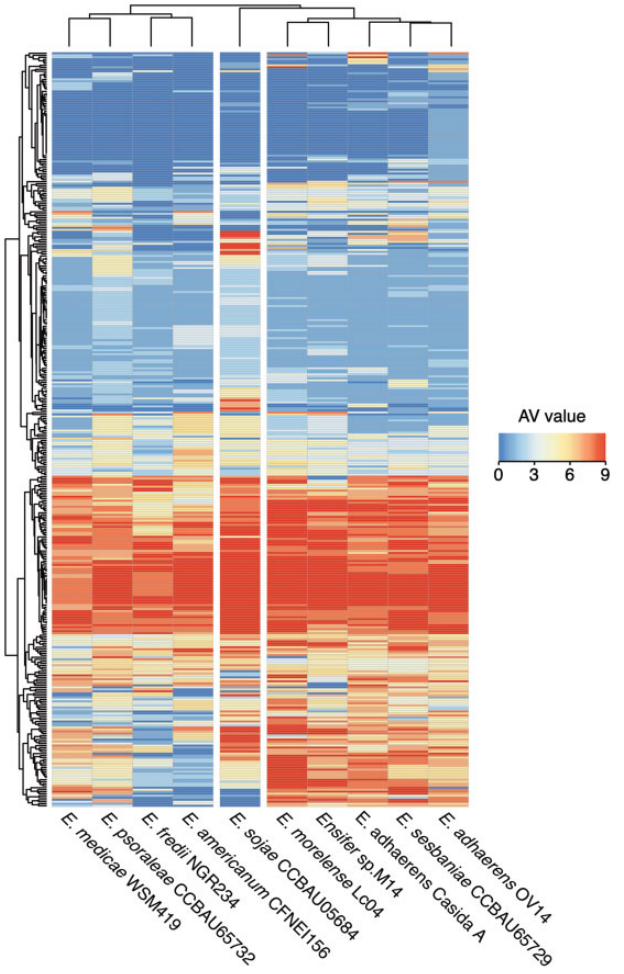


Fig. 5.—Phenotypic properties of phylogenetically diverse *Ensifer* strains. Ten *Ensifer* strains were screened for their ability to catabolize 190 carbon sources, and to grow in 96 osmolyte and 96 pH conditions using Biolog Phenotype MicroArray plates PM1, PM2, PM9, and PM10. Growth in each well was summarized on a scale of 0 (dark blue) through 9 (dark red), with higher numbers representing more robust growth. A larger version of this figure, in which each condition is labeled along the *Y* axis, is provided as supplementary figure S8, Supplementary Material online.

$0.44 \pm 0.08\,h^{-1}$; $P = 0.03$ from an ANOVA followed by Tukey's test; supplementary fig. S7*A* and table S5, Supplementary Material online). On the other hand, tested strains of the symbiotic clade were, on average, better able to withstand heat stress (37 °C) in TY media (fig. 4*B*). No statistically significant difference in the average ability of the tested strains of the two clades to grow in minimal media with succinate or glucose as a carbon source, or with *M. sativa* (a symbiotic partner of *E. meliloti* and *E. medicae*) root exudates as a nitrogen source, was detected (supplementary fig. S7 and table S5, Supplementary Material online).

The phenotypic properties of the genus *Ensifer* were further examined through evaluating the ability of the same ten strains to catabolize 190 carbon sources, and to grow in 96 osmolyte and 96 pH conditions, through the use of Biolog Phenotype MicroArrays. Clustering the strains based on growth properties largely separated the tested strains of the symbiotic clade and nonsymbiotic clade into distinct groups (fig. 5 and supplementary fig. S8, Supplementary Material online). The exception was *Ensifer sojae*, which formed its own intermediate group in the phenotype data. To aid in identifying which conditions best separate the tested strains of the symbiotic clade (including *E. sojae*) from the tested strains of the nonsymbiotic clade, a linear discriminant analysis (LDA) was run over the AV values summarizing growth in each condition (supplementary data set S5, Supplementary Material online). In general, tested strains from the nonsymbiotic clade better tolerated high pH (pH 9.0–9.5) than did the tested strains from the symbiotic cluster. In contrast, tested strains of the symbiotic clade had better tolerance to low pH conditions (pH 3.5–4.5). In addition, the tested strains from the two clades clearly differed in their overall metabolic abilities with tested strains of the nonsymbiotic clade generally having a broader metabolic capacity than those of the symbiotic clade (supplementary table S6, Supplementary Material online). Whereas tested strains of the symbiotic clade displayed robust growth on 65 carbon sources on average, the tested strains of the nonsymbiotic clade grew on an average of 81 carbon sources ($P < 0.05$, Student's *t*-test). Overall, these initial experiments provide support for the hypothesis that a variety of phenotypes, not just the ability to nodulate legumes, differ between the symbiotic and nonsymbiotic clades of the genus *Ensifer*.

Discussion

We investigated the evolution of SNF within the genus *Ensifer*, which includes a mix of nitrogen-fixing bacteria and bacteria unable to fix nitrogen, as a model for the evolution of interkingdom mutualisms. Our results indicate that, despite SNF having likely evolved multiple times within the genus *Ensifer*, the symbiotic and nonsymbiotic strains are largely separated into two phylogenetic clades reminiscent of the general division of pathogenic and environmental strains between the genera *Burkholderia* and *Paraburkholderia* (Sawana et al. 2014). Although it is possible that this result will fail to remain true as more genome sequences are published, we believe the result to be robust as the current set of genome sequences are of strains isolated from diverse legumes and other diverse environments including the rhizosphere, pristine caves, and an abandoned mine. In addition to the prevalence of SNF, the two clades differ with respect to their genomic composition (pangenome content and genome size) and phenotypic properties (metabolic capacity and stress tolerance) based on initial tests of a subset of strains. There have been several revisions to the taxonomy of the genus *Ensifer*. Recently, a

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

genome-based approach was proposed to standardize bacterial taxonomy (Parks et al. 2018) that splits the genus *Ensifer* into two genera: *Sinorhizobium* and *Ensifer*. The symbiotic and nonsymbiotic clades identified here correspond with the genera *Sinorhizobium* and *Ensifer*, respectively, supporting the proposal to divide the genus *Ensifer* into two genera.

Our analyses revealed a complex evolutionary history of SNF within the genus *Ensifer*. In addition to SNF emerging a predicted six times or more, we detected possible losses of SNF and allele switches. Between one and six of the NodABC and NifHDK proteins were not detected in five of the strains in the symbiotic clade (fig. 1). Although this may be indicative of multiple losses of symbiosis, we cannot rule out that these are false negatives due to incomplete genome assemblies or genome assembly errors; five of the six genomes were draft genomes, and the one strain with a complete genome (*E. meliloti* M162) can nodulate ten of 27 tested *Medicago truncatula* genotypes suggesting it does contain *nod* and *nif* genes (Sugawara et al. 2013). Based on the RepAB phylogeny (fig. 2), the symbiotic megaplasmid of *Ensifer arboris* likely shares common ancestry with the symbiotic megaplasmids of the sister species *E. meliloti* and *E. medicae*. Yet, the nodulation and nitrogen fixation genes appeared distinct. Thus, we hypothesize that there was a recent replacement of the symbiotic genes in *E. arboris*, or alternatively, in *E. meliloti* and *E. medicae*. This hypothesis is supported by the observation that the host ranges of these rhizobia differ; unlike *E. meliloti* and *E. medicae*, *E. arboris* does not nodulate plants of the genus *Medicago* (Zhang et al. 1991).

The reason for the phylogenetic bias in the evolution of SNF within the genus *Ensifer* remains unclear, especially considering that strains from both clades are plant-associated (Bai et al. 2015). One hypothesis is that each clade occupies distinct niches within the soil and plant-associated environments. Indeed, analysis of a subset of strains suggested that species of the nonsymbiotic clade have a broader metabolic capacity (supplementary table S6, Supplementary Material online), as well as a larger average genome size (fig. 3*B*), which is consistent with these species being more capable of adapting to fluctuating nutritional environments. This is further supported by the apparently higher number of rRNA operons in strains of the nonsymbiotic clade, which is generally thought to allow bacteria to more quickly respond to changing nutrient conditions (Stevenson and Schmidt 2004; Roller et al. 2016). Moreover, the nonsymbiotic clade could be differentiated from the symbiotic clade based on its pangenome content (fig. 3*A*), which leads us to hypothesize that strains of these clades acquire genes from distinct gene pools, further supporting the hypothesis that they belong to distinct gene-cohesive groups and ecological niches. This hypothesis may then be interpreted in the framework of the stable ecotype model (sensu Cohan 2006), where the symbiotic and nonsymbitic clades represent two, ecologically distinct and monophyletic groups and where periodic selection events (e.g., fitness for SNF) are recurrent.

An alternate, but not mutually exclusive, hypothesis is that the symbiotic *Ensifer* clade contains "facilitator" genes required to support SNF, similar to the theory that ancestral legumes contained a genetic "predisposition" necessary for the eventual evolution of rhizobium symbioses (Soltis et al. 1995; Doyle 2011; Werner et al. 2014). Conversely, evolution of SNF may also require the absence of "inhibitor" genes, such as the absence of virulence factors (Marchetti et al. 2010). As we did not evaluate cause-and-effect relationships, our data set does not definitely address these hypotheses. However, we observed numerous genotypic and likely phenotypic differences between the symbiotic and nonsymbiotic clade, providing some support for these hypotheses. For example, the tested strains of the symbiotic clade appeared to have higher tolerance to low pH (supplementary fig. S8 and data set S5, Supplementary Material online), which is notable as the curled root hair is an acidic environment (Hawkins et al. 2017). At the genomic level, 231 of the core genes of the symbiotic clade were absent from the pangenome of the nonsymbiotic clade and thus are good candidates as possible facilitators and follow-up studies. However, facilitators and inhibitors could also take the form of polymorphisms within highly conserved genes, as shown for *bacA* and the *Sinorhizobium–Medicago* symbiosis (diCenzo et al. 2017).

In summary, we show that the legume symbionts and nonsymbionts of the genus *Ensifer* are largely segregated into two phylogenetically distinct clades that differ in their genomic and phenotypic properties. We suggest that these observations, which follow the guidelines recently reported for rhizobia and agrobacteria (de Lajudie et al. 2019), support the division of the genus into two genera: *Ensifer* for the nonsymbiotic clade and *Sinorhizobium* for the symbiotic clade. However, formal descriptions and publication of the genera in the International Journal of Systematic and Evolutionary Microbiology (IJSEM) are still required. We also provide evidence that SNF genes were likely acquired several independent times within this genus, but predominately within one monophyletic clade. These observations suggest that the presence or absence of other genomic features ("facilitators" or "inhibitors") aside from the core symbiotic genes could be required for the establishment of an effective symbiosis. This suggestion is supported by the ability to differentiate the strains of the two clades based on their pangenome content and, at least for the tested subset, their phenotypic properties. However, as cause-and-effect relationships were not examined, follow-up study is required to more directly test this facilitators hypothesis.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

Acknowledgments

We are grateful to E. Martinez-Romero for providing strains *Ensifer morelense* Lc04 and *E. morelense* Lc18, to E. Mullins (Teagasc, MTA2018233) for *Ensifer adhaerens* OV14, to C.-F. Tian for *Sinorhizobium fredii* NGR 234, and to L. Dziewit for *Ensifer* sp. M14. A.M. was supported by the Fondazione Cassa di Risparmio di Firenze (Grant No. 18204, 2017.0719), by the "MICRO4Legumes" grant (Italian Ministry of Agriculture), and by the grant "Dipartimento di Eccellenza 2018–2022" from the Italian Ministry of Education, University and Research (MIUR). L.C. was supported by the MICRO4Legumes grant (Italian Ministry of Agriculture). G.C.D. was supported by a postdoctoral fellowship from the Natural Science and Engineering Research Council of Canada (NSERC), funding from Queen's University, and a NSERC Discovery Grant.

Data Availability

Scripts to repeat the computational analyses reported in this study are available at https://github.com/diCenzo-GC/Ensifer_phylogenomics (last accessed October 18,2020).

Literature Cited

Bai Y, et al. 2015. Functional overlap of the Arabidopsis leaf and root microbiota. Nature 528(7582):364–369.

Bankevich A, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19(5):455–477.

Barcellos FG, Menna P, Batista JS, da S, Hungria M. 2007. Evidence of horizontal transfer of symbiotic genes from a *Bradyrhizobium japonicum* inoculant strain to indigenous diazotrophs *Sinorhizobium* (*Ensifer*) *fredii* and *Bradyrhizobium elkanii* in a Brazilian savannah soil. Appl Environ Microbiol. 73(8):2635–2643.

Barzanti R, et al. 2007. Isolation and characterization of endophytic bacteria from the nickel hyperaccumulator plant *Alyssum bertolonii*. Microb Ecol. 53(2):306–316.

Biondi EG, et al. 2009. Metabolic capacity of *Sinorhizobium* (*Ensifer*) *meliloti* strains as determined by phenotype MicroArray analysis. Appl Environ Microbiol. 75(16):5396–5404.

Bochner B, Gomez V, Ziman M, Yang S, Brown SD. 2010. Phenotype microarray profiling of *Zymomonas mobilis* ZM4. Appl Biochem Biotechnol. 161(1–8):116–123.

Bosi E, et al. 2015. MeDuSa: a multi-draft based scaffolder. Bioinformatics 31(15):2443–2451.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25(15):1972–1973.

Casida LE. 1982. *Ensifer adhaerens* gen. nov., sp. nov.: a bacterial predator of bacteria in soil. Int J Syst Evol Microbiol. 32:339–345.

Checcucci A, et al. 2016. Mixed nodule infection in *Sinorhizobium meliloti–Medicago sativa* symbiosis suggest the presence of cheating behavior. Front Plant Sci. 7:835.

Checcucci A, et al. 2017. Role and regulation of ACC deaminase gene in *Sinorhizobium meliloti*: is it a symbiotic, rhizospheric or endophytic gene? Front Genet. 8:6.

Checcucci A, diCenzo GC, Perrin E, Bazzicalupo M, Mengoni A. 2019. Genomic diversity and evolution of rhizobia. In: Das S, Dash HR, editors. Microbial diversity in the genomic era. London, United Kingdom: Academic Press. p. 37–46.

Cohan FM. 2006. Towards a conceptual and operational union of bacterial systematics, ecology, and evolution. Philos Trans R Soc B. 361(1475):1985–1996.

Cowie A, et al. 2006. An integrated approach to functional genomics: construction of a novel reporter gene fusion library for *Sinorhizobium meliloti*. Appl Environ Microbiol. 72(11):7156–7167.

de Lajudie PM, et al. 2019. Minimal standards for the description of new genera and species of rhizobia and agrobacteria. Int J Syst Evol Microbiol. 69(7):1852–1863.

diCenzo GC, Tesi M, Pfau T, Mengoni A, Fondi M. 2020. Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. Nat Commun 11:2574.

diCenzo GC, et al. 2019. Multi-disciplinary approaches for studying rhizobium–legume symbioses. Can J Microbiol. 65(1):1–33.

diCenzo GC, Zamani M, Ludwig HN, Finan TM. 2017. Heterologous complementation reveals a specialized activity for BacA in the *Medicago–Sinorhizobium meliloti* symbiosis. Mol Plant Microbe Interact. 30(4):312–324.

Doin de Moura GG, Remigi P, Masson-Boivin C, Capela D. 2020. Experimental evolution of legume symbionts: what have we learnt? Genes 11(3):339.

Douglas AE. 2014. Symbiosis as a general principle in eukaryotic evolution. Cold Spring Harb Perspect Biol. 6(2):a016113.

Doyle JJ. 2011. Phylogenetic perspectives on the origins of nodulation. Mol Plant Microbe Interact. 24(11):1289–1295.

Drewniak L, Matlakowska R, Sklodowska A. 2008. Arsenite and arsenate metabolism of *Sinorhizobium* sp. M14 living in the extreme environment of the Zloty Stok gold mine. Geomicrobiol J. 25(7–8):363–370.

Dunkley EJ, Chalmers JD, Cho S, Finn TJ, Patrick WM. 2019. Assessment of phenotype microarray plates for rapid and high-throughput analysis of collateral sensitivity networks. PLoS One 14(12):e0219879.

Eddy SR. 2009. A new generation of homology search tools based on probabilistic inference. Genome Inform. 23:205–211.

Finan TM, Kunkel B, De Vos GF, Signer ER. 1986. Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. J Bacteriol. 167(1):66–72.

Finn RD, et al. 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 44(D1):D279–D285.

Galardini M, et al. 2014. DuctApe: a suite for the analysis and correlation of genomic and OmniLog^{TM} phenotype microarray data. Genomics 103(1):1–10.

Garrido-Oter R, et al. 2018. Modular traits of the Rhizobiales root microbiota and their evolutionary relationship with symbiotic rhizobia. Cell Host Microbe. 24(1):155–167.

Geddes BA, Kearsley J, Morton R, diCenzo GC, Finan TM. 2020. The genomes of rhizobia. In: Frendo P, Frugier F, Masson-Boivin C, editors. Regulation of nitrogen-fixing symbioses in legumes. Vol. 94. London, United Kingdom: Academic Press. p. 213–249.

Gerhart J, Kirschner M. 2007. The theory of facilitated variation. Proc Natl Acad Sci U S A. 104(Suppl 1):8582–8589.

Griesmann M, et al. 2018. Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. Science. 361(6398):eaat1743.

Haag AF, et al. 2013. Molecular insights into bacteroid development during *Rhizobium*–legume symbiosis. FEMS Microbiol Rev. 37(3):364–383.

Haft DH, et al. 2012. TIGRFAMs and genome properties in 2013. Nucleic Acid Res. 41(D1):D387–D395.

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

Haskett TL, et al. 2016. Assembly and transfer of tripartite integrative and conjugative genetic elements. Proc Natl Acad Sci U S A. 113(43):12268–12273.
Hawkins JP, Geddes BA, Oresnik IJ. 2017. Succinoglycan production contributes to acidic pH tolerance in *Sinorhizobium meliloti* Rm1021. Mol Plant Microbe Interact. 30(12):1009–1019.
Heath KD, Tiffin P. 2007. Context dependence in the coevolution of plant and rhizobial mutualists. Proc R Soc B. 274(1620):1905–1912.
Hooykaas PJ, Snijdewint FG, Schilperoort RA. 1982. Identification of the Sym plasmid of *Rhizobium leguminosarum* strain 1001 and its transfer to and expression in other rhizobia and *Agrobacterium tumefaciens*. Plasmid 8(1):73–82.
Howieson JG, Ewing MA. 1986. Acid tolerance in the *Rhizobium meliloti–Medicago symbiosis*. Aust J Agric Res. 37(1):55–64.
Hyatt D, et al. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11(1):119.
Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun. 9(1):7200.
Johnson DA, et al. 2008. High-throughput phenotypic characerization of *Pseudomonas aeruginosa* membrane transport genes. PLoS Genet. 4(10):e1000211.
Kalvari I, et al. 2018. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. Nucleic Acids Res. 46(D1):D335–D342.
Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.
Kolbe DL, Eddy SR. 2011. Fast filtering for RNA homology search. Bioinformatics 27(22):3102–3109.
Koren S, et al. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 27(5):722–736.
Laslett D, Canbäck B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res. 32(1):11–16.
Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 44(W1):W242–W245.
Li QQ, et al. 2011. *Ensifer sojae* sp. nov., isolated from root nodules of *Glycine max* grown in saline-alkaline soils. Int J Syst Evol Microbiol. 61(Pt 8):1981–1988.
Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22(13):1658–1659.
Long SR. 2001. Genes and signals in the rhizobium–legume symbiosis. Plant Physiol. 125(1):69–72.
López-García P, Eme L, Moreira D. 2017. Symbiosis in eukaryotic evolution. J Theor Biol. 434:20–33.
Löytynoja A. 2014. Phylogeny-aware alignment with PRANK. In: Russell D, editor. Methods in molecular biology (methods and protocols). Vol. 1079. Totowa (NJ): Humana Press. p. 155–170.
Marchetti M, et al. 2010. Experimental evolution of a plant pathogen into a legume symbiont. PLoS Biol. 8(1):e1000280.
Martens M, et al. 2007. Multilocus sequence analysis of *Ensifer* and related taxa. Int J Syst Evol Microbiol. 57(3):489–503.
Masson-Boivin C, Giraud E, Perret X, Batut J. 2009. Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? Trends Microbiol. 17(10):458–466.
Masson-Boivin C, Sachs JL. 2018. Symbiotic nitrogen fixation by rhizobia—the roots of a success story. Curr Opin Plant Biol. 44:7–15.
Nandasena KG, O'Hara GW, Tiwari RP, Sezmiş E, Howieson JG. 2007. *In situ* lateral transfer of symbiosis islands results in rapid evolution of diverse competitive strains of mesorhizobia suboptimal in symbiotic nitrogen fixation on the pasture legume *Biserrula pelecinus* L. Environ Microbiol. 9(10):2496–2511.
Oksanen J, et al. 2018. vegan: community ecology package. R package version 2.5-3. Available from: https://CRAN.R-project.org/package=vegan. Accessed October 18, 2020.
Oldroyd GED, Murray JD, Poole PS, Downie JA. 2011. The rules of engagement in the legume-rhizobial symbiosis. Annu Rev Genet. 45(1):119–144.
Ormeño-Orrillo E, et al. 2015. Taxonomy of rhizobia and agrobacteria from the *Rhizobiaceae* family in light of genomics. Syst Appl Microbiol. 38(4):287–291.
Page AJ, et al. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31(22):3691–3693.
Parks DH, et al. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol. 36(10):996–1004.
Pérez Carrascal OM, et al. 2016. Population genomics of the symbiotic plasmids of sympatric nitrogen-fixing *Rhizobium* species associated with *Phaseolus vulgaris*. Environ Microbiol. 18(8):2660–2676.
Perrin E, et al. 2015. Genomes analysis and bacteria identification: the use of overlapping genes as molecular markers. J Microbiol Methods. 117:108–112.
Pfau T, et al. 2018. The intertwined metabolism during symbiotic nitrogen fixation elucidated by metabolic modelling. Sci Rep. 8(1):12504.
Poole P, Ramachandran V, Terpolilli J. 2018. Rhizobia: from saprophytes to endosymbionts. Nat Rev Microbiol. 16(5):291–303.
Rinaudi LV, González JE. 2009. The low-molecular-weight fraction of exopolysaccharide II from *Sinorhizobium meliloti* is a crucial determinant of biofilm formation. J Bacteriol. 191(23):7216–7224.
Rogel MA, HernÃÂ¿Â½ndez-Lucas I, Kuykendall LD, Balkwill DL, Martinez-Romero E. 2001. Nitrogen-fixing nodules with *Ensifer adhaerens* harboring *Rhizobium tropici* symbiotic plasmids. Appl Environ Microbiol. 67(7):3264–3268.
Roller BRK, Stoddard SF, Schmidt TM. 2016. Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. Nat Microbiol. 1(11):16160.
Sanjuán J. 2016. Towards the minimal nitrogen-fixing symbiotic genome. Environ Microbiol. 18(8):2292–2294.
Sawana A, Adeolu M, Gupta RS. 2014. Molecular signatures and phylogenomic analysis of the genus *Burkholderia*: proposal for division of this genus into the emended genus *Burkholderia* containing pathogenic organisms and a new genus *Paraburkholderia* gen. nov. harboring environmental species. Front Genet. 5:429.
Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30(14):2068–2069.
Soltis DE, et al. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. Proc Natl Acad Sci U S A. 92(7):2647–2651.
Sørensen M, et al. 2019. The role of exploitation in the establishment of mutualistic microbial symbioses. FEMS Microbiol Lett. 366(12):fnz148.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9):1312–1313.
Stevenson BS, Schmidt TM. 2004. Life history implications of rRNA gene copy number in *Escherichia coli*. Appl Environ Microbiol. 70(11):6670–6677.
Sugawara M, et al. 2013. Comparative genomics of the core and accessory genomes of 48 *Sinorhizobium* strains comprising five genospecies. Genome Biol. 14(2):R17.
Sullivan JT, Patrick HN, Lowther WL, Scott DB, Ronson CW. 1995. Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. Proc Natl Acad Sci U S A. 92(19):8985–8989.
Tian CF, Young JPW. 2019. Genomics and evolution of rhizobia. In: Wang ET, Tian CF, Chen WF, Young JPW, Chen WX,

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

editors. Ecology and evolution of rhizobia: principles and applications. Singapore: Springer. p. 103–119.
Toledo I, Lloret L, Martínez-Romero E. 2003. *Sinorhizobium americanus* sp. nov., a new *Sinorhizobium* species nodulating native *Acacia* spp. in Mexico. Syst Appl Microbiol. 26(1):54–64.
Trinick MJ. 1980. Relationships amongst the fast-growing rhizobia of *Lablab purpureus*, *Leucaena leucocephala*, *Mimosa* spp., *Acacia farnesiana* and *Sesbania grandiflora* and their affinities with other rhizobial groups. J Appl Bacteriol. 49(1):39–53.
van Velzen R, et al. 2018. Comparative genomics of the nonlegume *Parasponia* reveals insights into evolution of nitrogen-fixing rhizobium symbioses. Proc Natl Acad Sci U S A. 115:E4700–E4709.
Vasilinetc I, Prjibelski AD, Gurevich A, Korobeynikov A, Pevzner PA. 2015. Assembling short reads from jumping libraries with large insert sizes. Bioinformatics 31(20):3262–3268.
Wang ET. 2019. Current systematics of rhizobia. In: Wang ET, Tian CF, Chen WF, Young JPW, Chen WX, editors. Ecology and evolution of rhizobia: principles and applications. Singapore: Springer. p. 41–102.
Wang ET, Romero JM, Romero EM. 1999. Genetic diversity of rhizobia from *Leucaena leucocephala* nodules in Mexican soils. Mol Ecol. 8(5):711–724.
Wang ET, Young JPW. 2019. History of rhizobial taxonomy. In: Wang ET, Tian CF, Chen WF, Young JPW, Chen WX, editors. Ecology and evolution of rhizobia: principles and applications. Singapore: Springer. p. 23–39.
Wang YC, et al. 2013. Proposal of *Ensifer psoraleae* sp. nov., *Ensifer sesbaniae* sp. nov., *Ensifer morelense* comb. nov. and *Ensifer americanum* comb. nov. Syst Appl Microbiol. 36(7):467–473.
Warnes GR, et al. 2016. gplots: various R programming tools for plotting data. R package version 3.0.1. Available from: https://CRAN.R-project.org/package=gplots. Accessed October 18, 2020.
Wendt T, Doohan F, Mullins E. 2012. Production of *Phytophthora infestans*-resistant potato (*Solanum tuberosum*) utilising *Ensifer adhaerens* OV14. Transgenic Res. 21(3):567–578.
Werner GDA, Cornwell WK, Cornelissen JHC, Kiers ET. 2015. Evolutionary signals of symbiotic persistence in the legume–rhizobia mutualism. Proc Natl Acad Sci U S A. 112(33):10262–10269.
Werner GDA, Cornwell WK, Sprent JI, Kattge J, Kiers ET. 2014. A single evolutionary innovation drives the deep evolution of symbiotic N_2-fixation in angiosperms. Nat Commun. 5(1):4087.
Wickham H. 2016. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag.
Willems A, et al. 2003. Description of new *Ensifer* strains from nodules and proposal to transfer *Ensifer adhaerens* Casida 1982 to *Sinorhizobium* as *Sinorhizobium adhaerens* comb. nov. Request for an opinion. Int J Syst Evol Microbiol. 53(4):1207–1217.
Wu M, Scott AJ. 2012. Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. Bioinformatics 28(7):1033–1034.
Young JM. 2003. The genus name *Ensifer* Casida 1982 takes priority over *Sinorhizobium* Chen et al. 1988, and *Sinorhizobium morelense* Wang et al. 2002 is a later synonym of *Ensifer adhaerens* Casida 1982. Is the combination '*Sinorhizobium adhaerens*' (Casida 1982) Willems et al. 2003 legitimate? Request for an opinion. Int J Syst Evol Microbiol. 53:2107–2110.
Zhang X, Harper R, Karisto M, Lindström K. 1991. Diversity of *Rhizobium* bacteria isolated from the root nodules of leguminous trees. Int J Syst Evol Microbiol. 41(1):104–113.
Zhao R, et al. 2018. Adaptive evolution of rhizobial symbiotic compatibility mediated by co-evolved insertion sequences. ISME J. 12(1):101–111.

Associate editor: Esperanza Martinez-Romero

Downloaded from https://academic.oup.com/gbe/article/12/12/2521/5923295 by Biblioteca di Scienze, Università degli studi di Firenze user on 31 January 2024

AMERICAN SOCIETY FOR MICROBIOLOGY | mSystems®

RESEARCH ARTICLE

DNA Methylation in *Ensifer* Species during Free-Living Growth and during Nitrogen-Fixing Symbiosis with *Medicago* spp.

George C. diCenzo,[a,b] Lisa Cangioli,[b] Quentin Nicoud,[c] Janis H. T. Cheng,[a] Matthew J. Blow,[d] Nicole Shapiro,[d] Tanja Woyke,[d] Emanuele G. Biondi,[c] Benoît Alunni,[c] Alessio Mengoni,[b] Peter Mergaert[c]

[a]Department of Biology, Queen's University, Kingston, Ontario, Canada
[b]Department of Biology, University of Florence, Florence, Italy
[c]Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell, Gif-sur-Yvette, France
[d]U.S. Department of Energy Joint Genome Institute, Berkeley, California, USA

Peter Mergaert and Alessio Mengoni contributed equally to this work. Author order was determined alphabetically.

ABSTRACT Methylation of specific DNA sequences is ubiquitous in bacteria and has known roles in immunity and regulation of cellular processes, such as the cell cycle. Here, we explored DNA methylation in bacteria of the genus *Ensifer*, including its potential role in regulating terminal differentiation during nitrogen-fixing symbiosis with legumes. Using single-molecule real-time sequencing, six genome-wide methylated motifs were identified across four *Ensifer* strains, five of which were strain-specific. Only the GANTC motif, recognized by the cell cycle-regulated CcrM methyltransferase, was methylated in all strains. In actively dividing cell cultures, methylation of GANTC motifs increased progressively from the *ori* to *ter* regions in each replicon, in agreement with a cell cycle-dependent regulation of CcrM. In contrast, there was near full genome-wide GANTC methylation in the early stage of symbiotic differentiation. This was followed by a moderate decrease in the overall extent of methylation and a progressive decrease in chromosomal GANTC methylation from the *ori* to *ter* regions in later stages of differentiation. Based on these observations, we suggest that CcrM activity is dysregulated and constitutive during terminal differentiation, which we hypothesize is a driving factor for endoreduplication of terminally differentiated bacteroids.

IMPORTANCE Nitrogen fixation by rhizobia in symbiosis with legumes is economically and ecologically important. The symbiosis can involve a complex bacterial transformation—terminal differentiation—that includes major shifts in the transcriptome and cell cycle. Epigenetic regulation is an important regulatory mechanism in diverse bacteria; however, the roles of DNA methylation in rhizobia and symbiotic nitrogen fixation have been poorly investigated. We show that aside from cell cycle regulation, DNA methyltransferases are unlikely to have conserved roles in the biology of bacteria of the genus *Ensifer*. However, we present evidence consistent with an interpretation that the cell cycle methyltransferase CcrM is dysregulated during symbiosis, which we hypothesize may be a key factor driving the cell cycle switch in terminal differentiation required for effective symbioses.

KEYWORDS rhizobia, symbiotic nitrogen fixation, DNA methylation, cell cycle regulation, CcrM

Methylation of genomic DNA is a pervasive phenomenon found in eukaryotes (1–3), archaea (4), and bacteria (4, 5). The biological roles of DNA methylation are most extensively studied in mammals, where it contributes to normal development and disease via its impact on gene expression (6). In bacteria, DNA methylation is best known for its role in restriction-modification (R-M) systems that are thought to provide defense against phage

Editor Davide Bulgarelli, University of Dundee

Copyright © 2022 diCenzo et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to George C. diCenzo, george.dicenzo@queensu.ca.

The authors declare no conflict of interest.

Received 2 September 2021
Accepted 9 December 2021
Published 4 January 2022

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

infection and limit horizontal gene transfer through the degradation of invading nonmethylated DNA (7). Several methyltransferases (MTases) of R-M systems have also been implicated in phase variation in pathogens through modulating gene expression (8). A recent study of over 200 bacterial and archaeal species identified orphan MTases not belonging to R-M systems in nearly half of the genomes (4). To date, biological functions have been attributed to very few orphan MTases, namely, the Dam MTase of the *Gammaproteobacteria* and the CcrM MTase of the *Alphaproteobacteria* (9). The Dam MTase of *Escherichia coli* is notable for its role in regulation of DNA replication (10, 11) and DNA repair (12) by modulating the activity of other DNA-binding proteins. The CcrM MTase was first identified in *Caulobacter crescentus* (13), with homologs since identified in diverse *Alphaproteobacteria* (14, 15). CcrM activity was shown to be cell cycle regulated in *C. crescentus* and *Agrobacterium tumefaciens* (13, 16), leading to methylation of its cognate DNA motif (the pentanucleotide GANTC) specifically during a short period at the end of DNA replication. This leads to a switching of GANTC sites between fully methylated (methylated on both strands) and hemi-methylated (methylated only on the template strand) as a result of DNA replication (17), which serves to modulate gene expression in a cell cycle-dependent fashion (18–20). Over- and underexpression of *ccrM* result in defects in DNA replication and cell division (14, 16, 19), while its complete loss is lethal under some conditions.

The rhizobia are a polyphyletic group of *Alphaproteobacteria* and *Betaproteobacteria* that can both live free in the soil and enter into an endosymbiotic interaction with legumes (21). This interaction begins following an exchange of signals between the free-living partners (22), and it culminates in the formation of a new organ known as a root nodule, within which the cytoplasm of plant cells contain thousands of N_2-fixing bacteria called bacteroids. Bacteroid formation results in the differential expression of more than a thousand genes (23, 24) and global changes in cellular metabolism (25). In legumes of the inverted-repeat lacking clade (IRLC) and the dalbergioid clade of the family *Papilionoideae*, bacteroid development involves an additional process of terminal differentiation (26, 27); in other legume clades, bacteroid differentiation is less pronounced and is reversible. Terminal bacteroid development, in contrast to reversible bacteroid formation, involves cell enlargement (bacteroids are 5- to 10-fold longer than their free-living counterparts) and genome endoreduplication (resulting in up to 24 copies of the genome per cell), indicative of a cell cycle transition occurring during differentiation (26). Indeed, the correct expression of cell cycle regulators in *Ensifer* (syn. *Sinorhizobium*) *meliloti*, a symbiont of *Medicago* species of the IRLC, is essential for the formation of functional bacteroids (28, 29), while overexpression of CcrM or disruption of the master cell cycle regulator CtrA can give rise to bacteroid-like morphology in free-living cells (14, 30). Additionally, mutants in the *E. meliloti* cell cycle regulators *divJ*, *cbrA*, and *cpdR1*, encoding three negative regulators of CtrA, form nonfunctional nodules in which bacteroids do not differentiate properly (28, 29, 31), and genes encoding several cell cycle regulators (including CcrM) are strongly downregulated in bacteroids (24). The differentiation and cell cycle switch of bacteroids is controlled by the legume host through the production of a large family of peptides, known as nodule-specific cysteine-rich (NCR) peptides (32–34).

Multiple studies have provided evidence that changes in the methylation status of the DNA of legume nodule cells contribute to symbiotic development (35–37). Conversely, it remains unknown if methylation of rhizobium DNA contributes to the regulation of N_2-fixation or bacteroid development. We are aware of only one study (38) comparing DNA methylation of a rhizobium (*Bradyrhizobium diazoefficiens* USDA110) between free-living and symbiotic states (soybean nodules) and comparing these changes with differential expression data. Intriguingly, the authors identified a DNA motif that was methylated specifically in bacteroids (38). However, no clear evidence was presented that methylation of this (or any other) motif is involved in transcriptional regulation, and the number of genes both differentially expressed and differentially methylated in bacteroids did not appear to be different than expected by chance. While these data may suggest that DNA methylation does not play a major role in regulating N_2-fixation by rhizobia, they do not address the role of DNA methylation in terminal bacteroid differentiation, as *B. diazoefficiens* undergoes reversible differentiation in soybean nodules (39).

Here, we use Pacific Biosciences single-molecule real-time (SMRT) sequencing to detect genome-wide patterns of DNA methylation in four strains belonging to the genus *Ensifer*.

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

TABLE 1 Methylated motifs identified in this study

Motif[a]	Type[b]	Count[c]	Frequency (motifs/kb)
E. meliloti 2011			
G**A**NTC CTN**A**G	m6A	11,169	1.67
RCG**C**CTC YGCGGAG	m4C	3,943	0.59
CGC**A**(N5)GTG GCGT(N5)C**A**C	m6A	1,085	0.16
E. meliloti FSM-MA			
G**A**NTC CTN**A**G	m6A	11,215	1.67
TCG**A**(N8)TCGA AGCT(N8)**A**GCT	m6A	2,612	0.39
E. fredii NGR234			
G**A**NTC CTN**A**G	m6A	11,111	1.61
CAG**A**(N7)GTTG GTCT(N7)C**A**AC	m6A	188	0.03
E. adhaerens OV14			
G**A**NTC CTN**A**G	m6A	8,475	1.10
WNCCG**A**TG WNGGCTAC	m6A	4,596	0.60

[a]The methylated nucleotides are indicated in boldface font.
[b]Indicates whether the modification is an N^6-methyladenoside (m6A) or N^4-methylcytosine (m4C).
[c]The total times the motif appears in the genome, regardless of methylation status.

Our results indicate that DNA methylation is poorly conserved across the genus, and interestingly, they reveal that the pattern of GANTC methylation in N_2-fixing bacteroids differs from that of free-living cells. Based on these data, we hypothesize that constitutive activation of the CcrM MTase may be a contributing factor driving terminal differentiation.

RESULTS

The methylomes of the genus *Ensifer*. Our experimental design, as summarized in Materials and Methods and Fig. S1 (at doi.org/10.6084/m9.figshare.16556205), was developed to support an investigation into multiple potential roles of DNA methylation in plant-associated bacteria from the genus *Ensifer* through the use of SMRT sequencing. This was accomplished by (i) including DNA samples isolated from phylogenetically diverse wild-type strains, (ii) examining DNA methylation in a single strain across multiple conditions (exponential-phase growth versus stationary phase; growth with sucrose versus growth with succinate), (iii) investigating the impact of a large-scale genome reduction on DNA methylation patterns, and (iv) isolating DNA from bacteroids purified from legume nodules.

We began with base modification analyses of four wild-type strains from three species, including three nodule-forming strains (*E. meliloti* Rm2011, *E. meliloti* FSM-MA, *Ensifer fredii* NGR234) and one plant-associated, nonsymbiotic strain (*Ensifer adhaerens* OV14). To ensure consistency, all strains were grown to mid-exponential phase in a common minimal medium with succinate as the carbon source. A total of six methylated motifs were identified, of which five were m6A modifications and one was a m4C modification (Table 1). Five of the six motifs were methylated specifically in one strain. Only the GANTC motif, recognized by the highly conserved cell cycle-regulated CcrM methyltransferase (13, 14), was methylated in all four strains. To further examine the conservation, or lack thereof, of DNA modification across the genus *Ensifer*, we examined the distribution of methyltransferases in the model species *E. meliloti*. Based on gene annotations, we identified 24 genes encoding putative MTases in a previous pangenome analysis of 20 *E. meliloti* strains (Table S1 at doi.org/10.6084/m9.figshare.16556205) (40). Of these 24 genes, only 1 (*ccrM*) was found in all 20 strains, while 4 were found in 2 strains and 19 were found in a single strain. These results

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

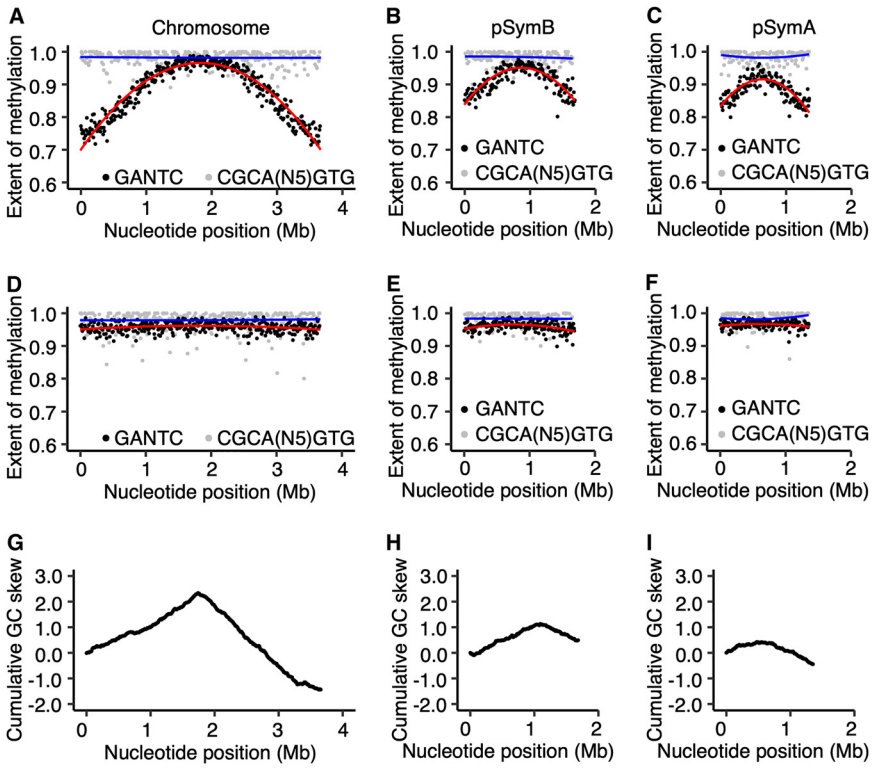


FIG 1 Genome-wide DNA methylation of *E. meliloti* Rm2011. (A to F) The extent of methylation is shown, using a 10-kb sliding window, of GANTC sites (black) and CGCA(N_5)GTG sites (gray) across the chromosome (A and D), pSymB (B and E), and pSymA (C and F) replicons of exponential phase (A to C) or early stationary phase (D to F) *E. meliloti* Rm2011. Averages from three biological replicates are shown. The red (GANTC) and blue [CGCA(N_5)GTG] lines are polynomial regression lines calculated in R using the "rlm" method and the formula "y~poly(x,2)." (G and I) Cumulative GC skews, shown using a 10-kb sliding window, across the *E. meliloti* Rm2011 chromosome (G), pSymB (H), and pSymA (I) replicons.

suggest that DNA methylation is unlikely to play a conserved role in the genus *Ensifer* aside from cell cycle control via CcrM-mediated methylation and phage defense.

None of the motifs methylated in *E. meliloti* Rm2011 were enriched in the promoter regions of genes previously shown to be differentially expressed when grown with glucose versus succinate (41). Similarly, except for the GANTC motif as discussed below, no global effect of carbon source (sucrose [glycolytic] versus succinate [gluconeogenic]) was observed on the DNA methylation pattern of *E. meliloti* Rm2011 (Fig. S2 at doi.org/10.6084/m9.figshare.16556205). Moreover, no global differences in DNA methylation were detected between *E. meliloti* Rm2011 and RmP3496, an Rm2011 derivative lacking the pSymA and pSymB replicons that together account for 45% of the genome content of *E. meliloti* (42) (Fig. S3 at doi.org/10.6084/m9.figshare.16556205). These results suggest that at least under the tested conditions, most DNA MTases are not likely to have a regulatory function in the genus *Ensifer*, although they could play regulatory roles in other environments.

Cell cycle regulation by the CcrM methyltransferase. A progressive increase in the extent of methylation (here defined as the estimated fraction of reads mapping to a motif that were methylated) of GANTC sites was observed from the *ori* to *ter* regions of the chromosomes of all four strains during mid-exponential growth (Fig. 1, and Fig. S4 to S6 at doi.org/10.6084/m9.figshare.16556205). There was a local drop in GANTC methylation around the 1.5-Mb mark in the *E. meliloti* FSM-MA chromosome (Fig. S4 at doi.org/10.6084/m9.figshare.16556205); however, this was seen in only two of three replicates and corresponded to a region of high sequencing depth (Fig. S7 at doi.org/10.6084/m9.figshare.16556205), suggesting the result is a sequencing artifact. In contrast to exponential-phase cultures, GANTC sites displayed near full methylation (averaging ~95%) across the genome during early stationary phase, while all other motifs displayed near full methylation (averaging 95 to 99%) across

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

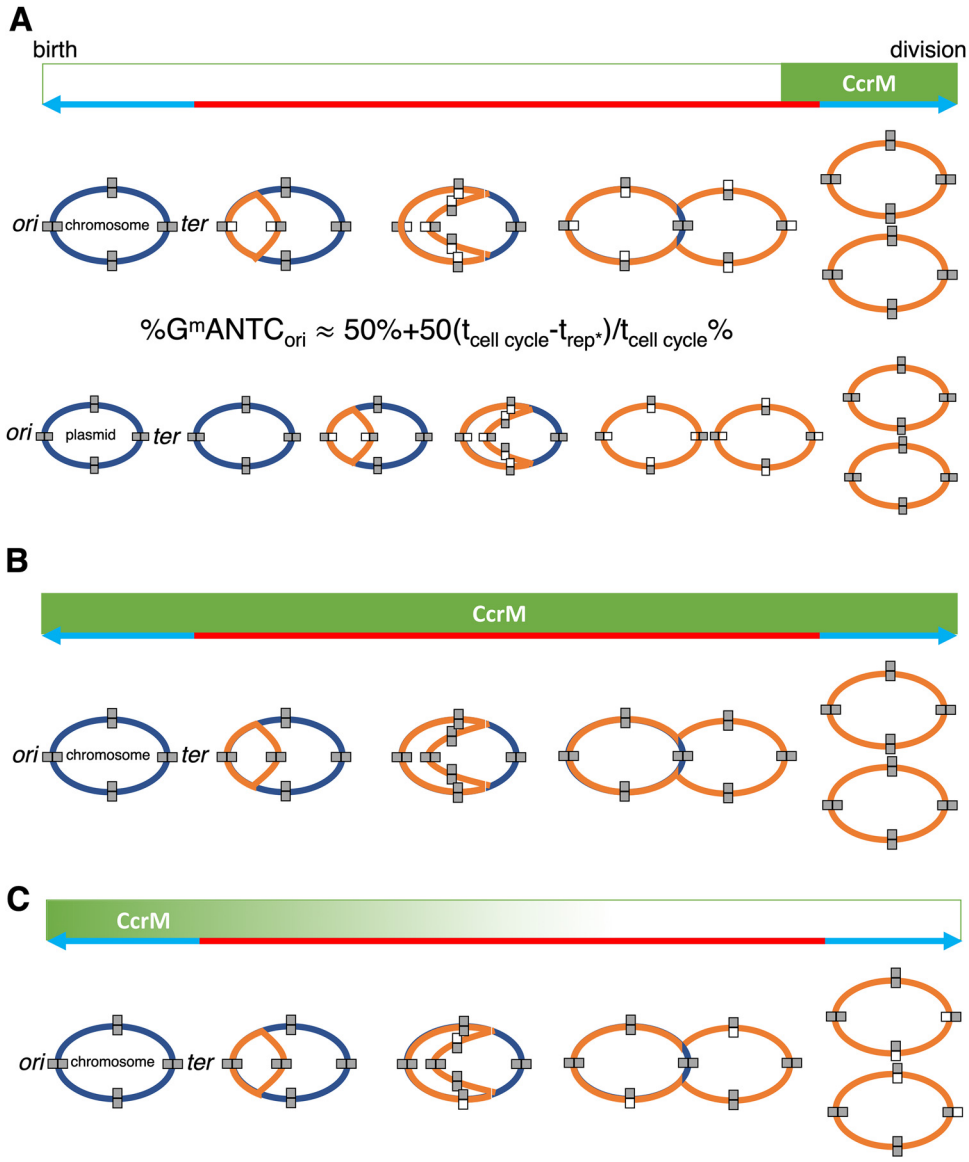


FIG 2 Model describing the GANTC methylation patterns observed in cultures and bacteroids. In all panels, the replication progression and methylation status of the circular chromosome and a megaplasmid are depicted during the cell cycle progression from cell birth until division. The replicons at cell birth are in dark blue, and the newly replicated DNA is in orange. Full-gray rectangles indicate GANTC sites that are in the fully methylated state, while gray/white rectangles indicate the hemimethylated state. The green boxes along the top indicate when CcrM is known or postulated to be active during the cell cycle, while the blue/red line indicates the relevant cell cycle phases; red is the S-phase or genome replication phase and blue is the gap phases and division. (A) The activity of CcrM in free-living cells during exponential growth. The formula expresses the extent of methylation at the origin (*ori*) of replication in an asynchronous bacterial population in culture. (B) The proposed activity of CcrM at an early stage of bacteroid differentiation, during which we suggest that the activity of CcrM is extended to also include the replication phase of the cell cycle. (C) The proposed activity of CcrM at a late stage of bacteroid differentiation, during which we suggest that there is a drop in CcrM activity during the last chromosome replication cycle of the endoreduplication process. The iconography of the illustrations is based on Fig. 1 of Mohapatra et al. (43). Refer to Fig. S22 at doi.org/10.6084/m9.figshare.16556205 for a version of this figure with a more detailed legend.

the genome regardless of growth state (Fig. 1, and Fig. S4 to S6 at doi.org/10.6084/m9.figshare.16556205). The observed pattern of GANTC methylation indicates a progressive switch from fully to hemi-methylated states as DNA replication proceeds (model provided as Fig. 2), confirming that the CcrM methyltransferase of the family *Rhizobiaceae* is cell cycle regulated as demonstrated in *C. crescentus* (13, 17, 43). Interestingly, the genome-wide pattern of GANTC methylation displayed a smaller variation in the extent of methylation from

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

the *ori* to *ter* regions in *E. meliloti* Rm2011 when grown with sucrose compared to succinate as the carbon source (Fig. S2 at doi.org/10.6084/m9.figshare.16556205). While this observation could suggest metabolic regulation of CcrM activity, we instead hypothesize, as displayed in Fig. 2A, that it is due to DNA replication being initiated later in the cell cycle when *E. meliloti* is provided sucrose, as recent observations showed that central carbon metabolism influences the rate of DNA polymerase processivity and timing of DNA replication initiation in *Bacillus subtilis* (44).

Notably, the GANTC methylation pattern differed across replicons within each genome. For example, in *E. meliloti* Rm2011 the extent of GANTC methylation ranged from 0.80 to 0.98 for pSymB and 0.80 to 0.96 for pSymA, while for the chromosome, the range was from 0.71 to 0.99 (Fig. 1). This result suggests that replication of each replicon is asynchronous, with replication of the secondary replicons being initiated later in the cell cycle than that of the chromosome (model provided as Fig. 2).

A previous study identified 462 cell cycle-regulated genes in *E. meliloti* Rm1021 (a near-isogenic relative of strain Rm2011 also derived from the nodule isolate SU47) through RNA sequencing of synchronized cell populations (45), which were classified into six groups based on the timing of their expression. We identified 111 cell cycle-regulated genes, belonging to 78 transcripts, that contained at least 1 GANTC in the predicted promoter regions (defined as the 125 bp upstream of the transcript; Data set S1 at doi.org/10.6084/m9.figshare.16556205), and distribution of these 111 genes across the 6 cell cycle gene expression groups was unbiased (45). As these 111 genes are both cell cycle-regulated and contain a GANTC site, they represent an initial candidate CcrM regulon in *E. meliloti*, although further work is required to identify the CcrM regulon.

Between 31 and 53 GANTC sites were repeatedly not detected as methylated on one or both strands in each of the four wild-type strains (see Data set S2 at doi.org/10.6084/m9.figshare.16556205), as has also been observed in other species (17, 46). However, further work is required to determine if this hypomethylation is biologically meaningful (e.g., CcrM cannot access the site due to binding of another protein [46]) or if it reflects technical artifacts (e.g., low sequencing coverage).

We found it striking that the *E. adhaerens* OV14 genomes had 2,636 to 2,740 fewer GANTC sites than the other three strains, despite having the largest genome size. Normalized by genome length, there are 1.10 GANTC sites per kb in the *E. adhaerens* OV14 genome (Table 1), which is similar to the 1.12 GANTC sites per kb in *C. crescentus* NA1000. In contrast, the three legume symbionts contained more than 1.60 GANTC sites per kb across their genomes (Table 1). This result prompted us to examine the frequency and distribution of GANTC sites across 157 *Ensifer* genomes. As defined previously (47), the genus *Ensifer* can be broadly subdivided into two monophyletic clades, the "symbiotic" clade (113 strains), in which nearly all strains are legume symbionts, and the "nonsymbiotic" clade (44 strains), in which nearly all strains are nonsymbionts (Fig. 3A). Consistent with previous results (18), GANTC sites occurred less frequently in all genomes (0.90 to 1.83 GANTC sites per kb) than expected in a random sequence of nucleotides (~3.5 GANTC sites per kb). Moreover, GANTC sites were ~2-fold more common in intergenic regions than in coding regions (Fig. 3B). Strikingly, there was a strong and statistically significant difference (P value $< 1 \times 10^{-10}$; two-sample t test) in the frequency of GANTC sites across the genomes of strains belonging to the symbiotic clade compared to strains of the nonsymbiotic clades (Fig. 3B), with an overall average of 1.70 and 1.06 GANTC sites per kb in the symbiotic and nonsymbiotic clades, respectively. The difference in the frequency of GANTC sites between the two clades could not be explained by differences in the GC content of these organisms, as both clades had an average GC content of 61.9% (Fig. S8 at doi.org/10.6084/m9.figshare.16556205), suggesting that the difference reflects underlying differences in the evolution, and possibly the biology, of these two clades.

The impact of bacteroid differentiation on DNA methylation. The only previously published study to examine the role of rhizobium DNA methylation during symbiosis using SMRT sequencing did so in a symbiosis where the bacteria do not undergo terminal differentiation (38, 39). To evaluate whether DNA methylation potentially contributes to

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

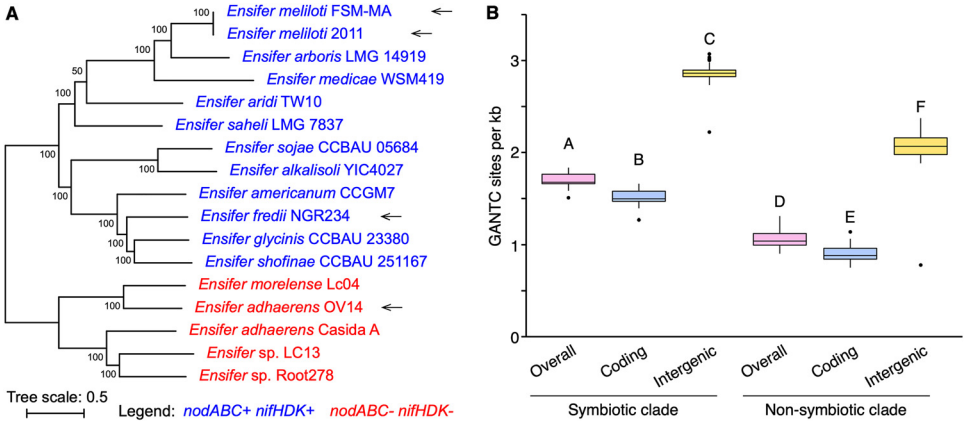


FIG 3 GANTC frequency in the genus *Ensifer*. (A) An unrooted maximum likelihood phylogeny of 17 representative *Ensifer* strains. The phylogeny represents the bootstrap best tree following 100 bootstrap replicates, prepared on the basis of the concatenated nucleotide alignments of 1,566 core genes. Values represent the bootstrap support. N_2-fixing legume symbionts were identified by the presence of the symbiotic genes *nodABC* and *nifHDK*. They are indicated in blue, while red denotes nonsymbiotic strains. The four wild-type strains used in this study are indicated with arrows. (B) Box plots summarizing the frequency of GANTC sites (presented as GANTC sites per kb) in 157 *Ensifer* strains is shown. The monophyletic "symbiotic" and "non-symbiotic" clades, as defined previously (47), are represented by 113 and 44 genomes, respectively. The densities of GANTC sites across the entire genome (pink), within coding regions (blue), and within intergenic regions (yellow) are shown. Statistically different values ($P < 0.05$) are denoted by uppercase letters as determined by a one-way analysis of variance (ANOVA) followed by Tukey's honestly significant difference (HSD) *post hoc* test.

regulation of terminal differentiation, we determined the DNA methylation patterns of *E. meliloti* Rm2011 and *E. meliloti* FSM-MA bacteroids purified from whole *Medicago sativa* nodules. *E. meliloti* FSM-MA bacteroids were additionally purified from *Medicago truncatula* nodules to determine whether the host plant influences bacteroid DNA methylation. *E. meliloti* Rm2011 bacteroids were not isolated from *M. truncatula* nodules as, unlike FSM-MA, Rm2011 forms a poor symbiosis with *M. truncatula* and produces moderately differentiated bacteroids in this host, whereas both strains are equally efficient on *M. sativa* (48, 49). See Kazmierczak et al. for a detailed description of the symbiotic phenotypes of these two strains of *E. meliloti* (49).

Moreover, we exploited the spatially distinct developmental zones that are present in indeterminate nodules (50), like those formed by *M. sativa* and *M. truncatula*. At the tip of these nodules, a bacterium-free meristem is present, responsible for the continuous growth of the nodule. Immediately below is the infection and differentiation zone II, where nodule cells become infected and bacteria differentiate into large, polyploid bacteroids. The tip and zone II of nodules appears white. Adjacent to the white zone II is the easily recognizable pinkish zone III (due to the presence of the oxygen-carrying leghemoglobin), where mature bacteroids fix nitrogen. This nodule tissue organization provided an opportunity to examine how DNA methylation patterns may differ between differentiating and differentiated bacteroids. To this end, *E. meliloti* Rm2011 and FSM-MA bacteroids were isolated from nodules hand-sectioned at the white-pink border; bacteroids isolated from the white sections represent the infecting and differentiating bacteroids (zone II), while those isolated from the pink sections represent the mature, hence terminally differentiated, N_2-fixing bacteroids (zone III) (see Fig. S9 at doi.org/10.6084/m9.figshare.16556205 for photographs of the plants).

Fluorescence microscopy and flow cytometry confirmed that nodule sectioning resulted in the isolation of distinct bacteroid populations (Fig. S10 to S13 at doi.org/10.6084/m9.figshare.16556205). Nearly all of the bacteroids isolated from whole-nodule samples and zone III samples were enlarged and polyploid, and most were positive for propidium iodide (PI) staining as expected for terminally differentiated bacteroids (26). In contrast, bacteroids of the zone II samples contained a mix of cell types differing in their size, ploidy level, and PI staining. These data confirmed that the whole-nodule samples and zone III samples consisted predominately of mature N_2-fixing bacteroids,

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

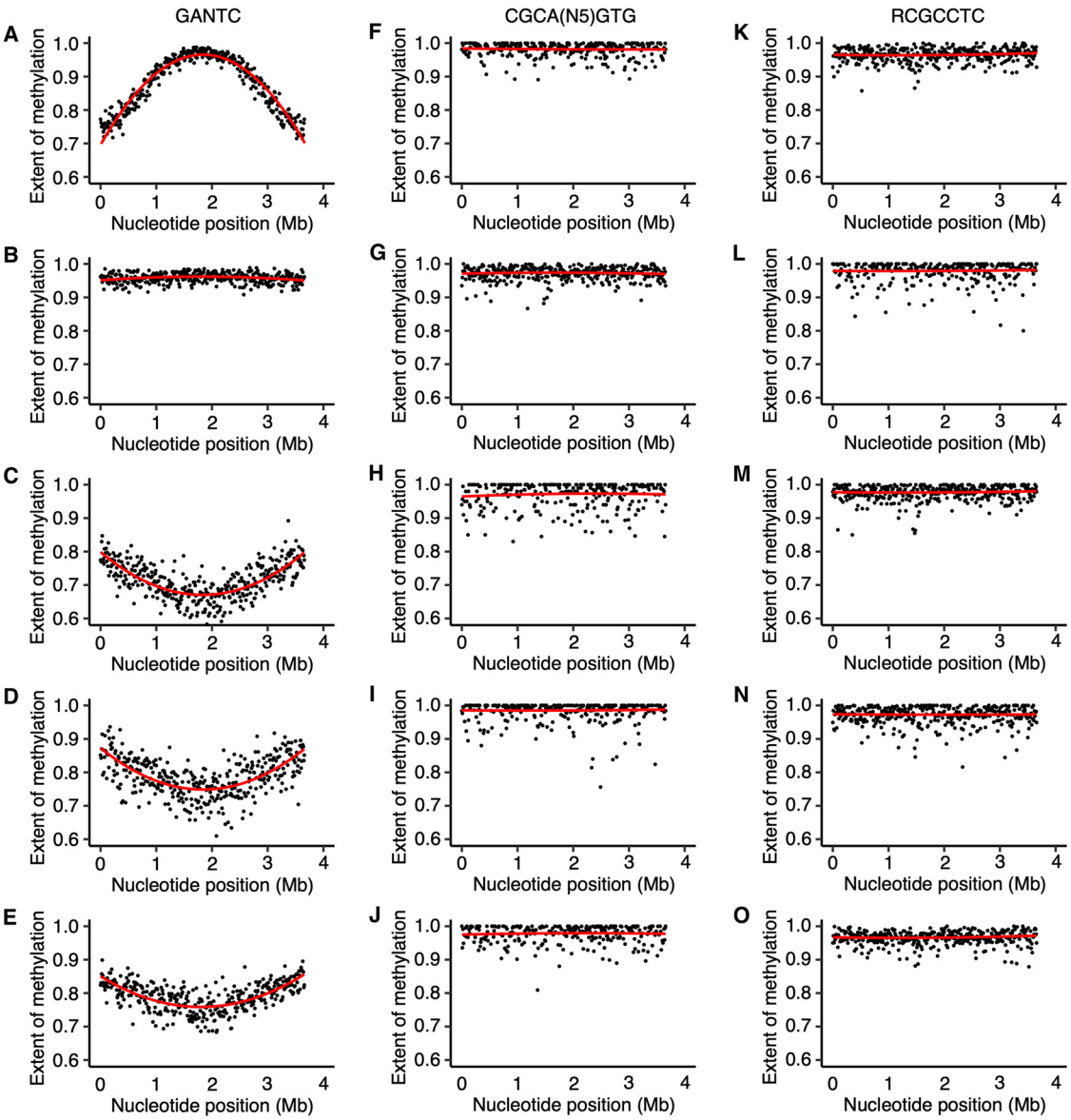


FIG 4 Chromosome-wide DNA methylation of *E. meliloti* Rm2011 bacteroids. (A to O) The extent of methylation of (A to E) GANTC, (F to J) CGCA(N_5)GTG, and (K to O) RCGCCTC motifs across the *E. meliloti* Rm2011 chromosome is shown using a 10-kb sliding window. Averages from three biological replicates are shown for free-living and whole-nodule samples; the data represent one replicate for the zone II and zone III nodule sections. (A, F, and K) Free-living cells harvested in mid-exponential phase. (B, G, and L) Free-living cells harvested in early stationary phase. (C, H, and M) Bacteroids isolated from *M. sativa* zone II nodule sections. (D, I, and N) Bacteroids isolated from *M. sativa* zone III nodule sections. (E, J, and O) Bacteroids isolated from *M. sativa* whole-nodule samples. The red lines are polynomial regression lines calculated in R using the "rlm" method and the formula "y~poly(x,2)." Data for pSymB and pSymA are shown in Fig. S14 and S15 at doi.org/10.6084/m9.figshare.16556205.

whereas the zone II samples contained a mix of cells at various stages of bacteroid differentiation.

With the exception of the GANTC sites (described below), no global difference was observed in the methylation patterns of bacteroids versus free-living cells (Fig. 4 and 5, and Fig. S14 to S17 at doi.org/10.6084/m9.figshare.16556205). Although a lower percentage of each motif was detected as methylated in the bacteroid samples compared to free-living samples, this was correlated with lower sequencing depth (Table S2 at doi.org/10.6084/m9.figshare.16556205) and thus unlikely to be biologically meaningful. Unlike *B. diazoefficiens* USDA110 (38), we did not identify motifs that were methylated specifically in the *E. meliloti* bacteroids. In addition, we found little evidence for any of the methylated motifs being enriched in the promoter regions of *E. meliloti* Rm2011 genes upregulated or

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

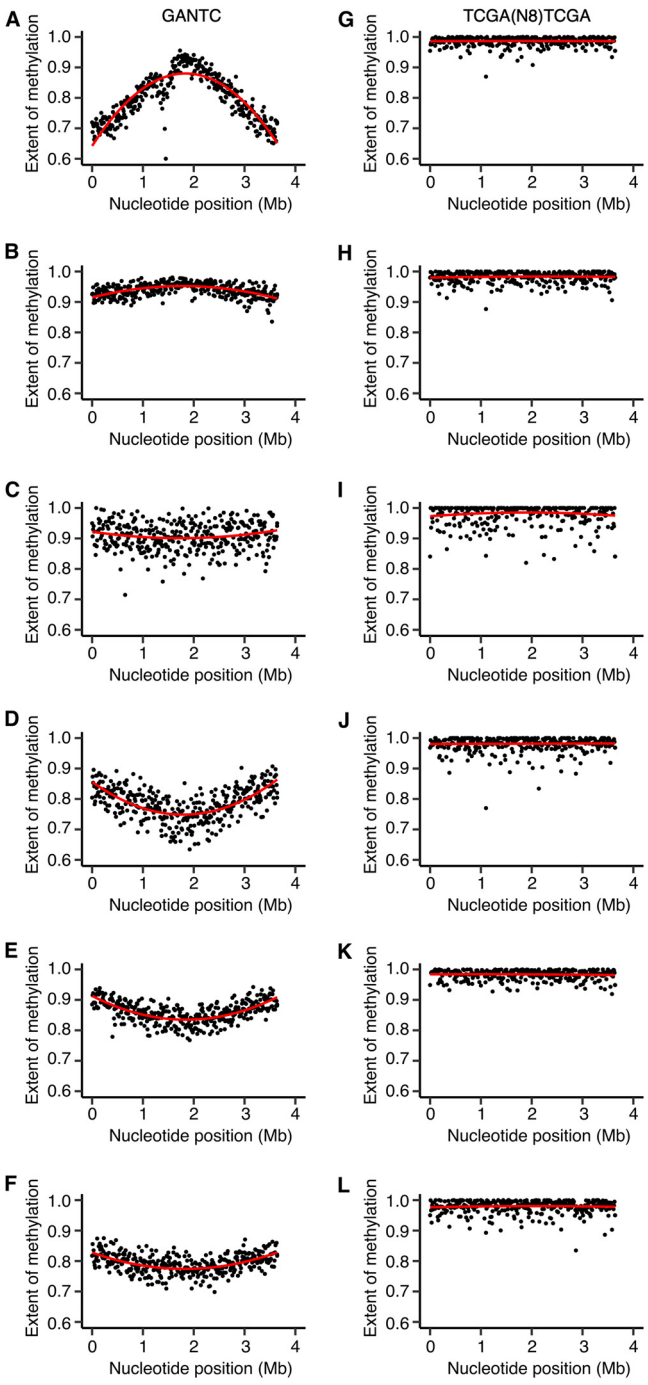


FIG 5 Chromosome-wide DNA methylation of *E. meliloti* FSM-MA bacteroids. (A to L) The extent of methylation of (A to F) GANTC and (G to L) TCGA(N_8)TCGA motifs across the *E. meliloti* FSM-MA chromosome is shown using a 10-kb sliding window. Averages from three biological replicates are shown for free-living and whole-nodule samples; the data represent one replicate for the zone II and zone III nodule sections. (A and G) Free-living cells harvested in mid-exponential phase. (B and H) Free-living cells harvested in early stationary phase. (C and I) Bacteroids isolated from *M. sativa* zone II nodule sections. (D and J) Bacteroids isolated from *M. sativa* zone III nodule sections. (E and K) Bacteroids isolated from *M. sativa* whole-nodule samples. (F and L) Bacteroids isolated from *M. truncatula* whole-nodule samples. The red lines are polynomial regression lines calculated in R using the "rlm" method and the formula "y~poly(x,2)." Data for pSymB and pSymA are shown in Fig. S16 and S17 at doi.org/10.6084/m9.figshare.16556205.

downregulated in bacteroids relative to free-living cells, as identified in published transcriptomic data for *E. meliloti* Rm1021 (23). These data suggest that most DNA methylation is unlikely to be a significant factor in directly regulating gene expression in *E. meliloti* bacteroids.

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

Bacteroid development influences GANTC methylation. Bacteroid development involves cell enlargement and genome endoreduplication, indicative of a cell cycle transition occurring during differentiation (26). Indeed, expression of *ccrM* and other genes encoding cell cycle regulators vary across stages of bacteroid development and are strongly downregulated in mature nitrogen-fixing *E. meliloti* bacteroids (24). We were therefore interested in examining whether GANTC methylation by the CcrM MTase was disrupted in bacteroids. Our data revealed a surprising genome-wide pattern of GANTC methylation in *E. meliloti* Rm2011 and FSM-MA bacteroids, which differed from free-living cells in either the exponential or stationary phases of growth (Fig. 4 and 5 and Fig. S14 to S17 at doi.org/10.6084/m9.figshare.16556205). The majority of GANTC sites had moderate to high levels of methylation in zone II, zone III, or whole-nodule samples, averaging 0.71 to 0.95 across each replicon (Fig. 4 and 5 and Fig. S14 to S17 and Tables S3 and S4 at doi.org/10.6084/m9.figshare.16556205). Most distinctively, a progressive decrease in the extent of chromosomal methylation of the GANTC sites was observed from *ori* to *ter* in nearly all bacteroid samples, revealing a characteristic smiling pattern, which differs from the patterns seen in exponential-phase (frowning pattern, i.e., a progressive increase from *ori* to *ter*) and stationary-phase (consistent methylation) cells. The exception was the *E. meliloti* FSM-MA zone II bacteroid sample, which displayed a consistently high level of GANTC methylation across the genome (Fig. 5C). This pattern, which is different from those in exponential-phase cells as well as mature bacteroids, could correspond to the methylation status of an early stage of bacteroid differentiation. We did not observe the same pattern in the *E. meliloti* Rm2011 zone II samples. As noted earlier, the zone II samples contain cells at various stages of differentiation (Fig. S10 to S13 at doi.org/10.6084/m9.figshare.16556205). Given that terminal differentiation is associated with an up to 24-fold increase in DNA content, small increases in the proportion of cells at late stages of differentiation could mask the DNA methylation pattern of the cells at early stages of differentiation. Thus, we hypothesize that the Rm2011 zone II sample captures a later stage of differentiation than that captured by the FSM-MA zone II sample. Supporting this hypothesis, the distribution of DNA content per cell in the flow cytometry data was flatter for *E. meliloti* Rm2011 zone II bacteroids compared to *E. meliloti* FSM-MA zone II bacteroids (Fig. S18 at doi.org/10.6084/m9.figshare.16556205), which suggests that the former sample represents a broader range of differentiation stages than the latter sample. In contrast to GANTC, the extent of methylation of the second m6A modified motif in each genome was consistently high, irrespective of condition or replicon (Fig. 4 and 5 and Fig. S14 to S17 and Tables S3 and S4 at doi.org/10.6084/m9.figshare.16556205). Similarly, sequencing depth was consistent across the length of each replicon (Fig. S7 at doi.org/10.6084/m9.figshare.16556205). These observations indicate that the changes in GANTC methylation patterns are biologically meaningful and not simply a sequencing artifact.

To further explore changes in CcrM activity during bacteroid differentiation, we took advantage of a collection of *M. truncatula* mutant plant lines (*dnf1*, *dnf2*, *dnf4*, *dnf5*, *dnf7*) whose nodules contain bacteria blocked at various stages of differentiation (see Fig. S19 at doi.org/10.6084/m9.figshare.16556205 for photographs of the plants) (51–57). Microscopy and flow cytometry data were consistent with past work (57) showing that bacteroids were blocked at the earliest to latest stages of differentiation in mutant plant lines in the order *dnf1* → *dnf5* → *dnf2* → *dnf7* → *dnf4* (Fig. 6). Nodule bacteria of *M. truncatula dnf1* mutant plants were small, with one or two haploid genome copies per cell (i.e., ploidy level = 1 or 2) (Fig. 6A and G), suggesting that the cell population was dominated by actively dividing cells that had not yet begun differentiation. Indeed, the GANTC methylation pattern of these cells (Fig. 6M and Fig. S20 and S21 at doi.org/10.6084/m9.figshare.16556205) resembled the frowning pattern of exponentially growing free-living cells (Fig. 5A). Although the nodule bacteria of *M. truncatula dnf5* mutant plants were also small and undifferentiated into bacteroids, the majority of cells had a ploidy level of 1 (Fig. 6B and H), suggesting these cells had ceased replication but had not yet begun the process of endoreduplication. GANTC methylation was consistently high across the chromosome of bacteria purified from *dnf5* nodules (Fig. 6N), similar to stationary-phase free-living cells (Fig. 6B) and indicating that terminal differentiation is preceded by full GANTC methylation.

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

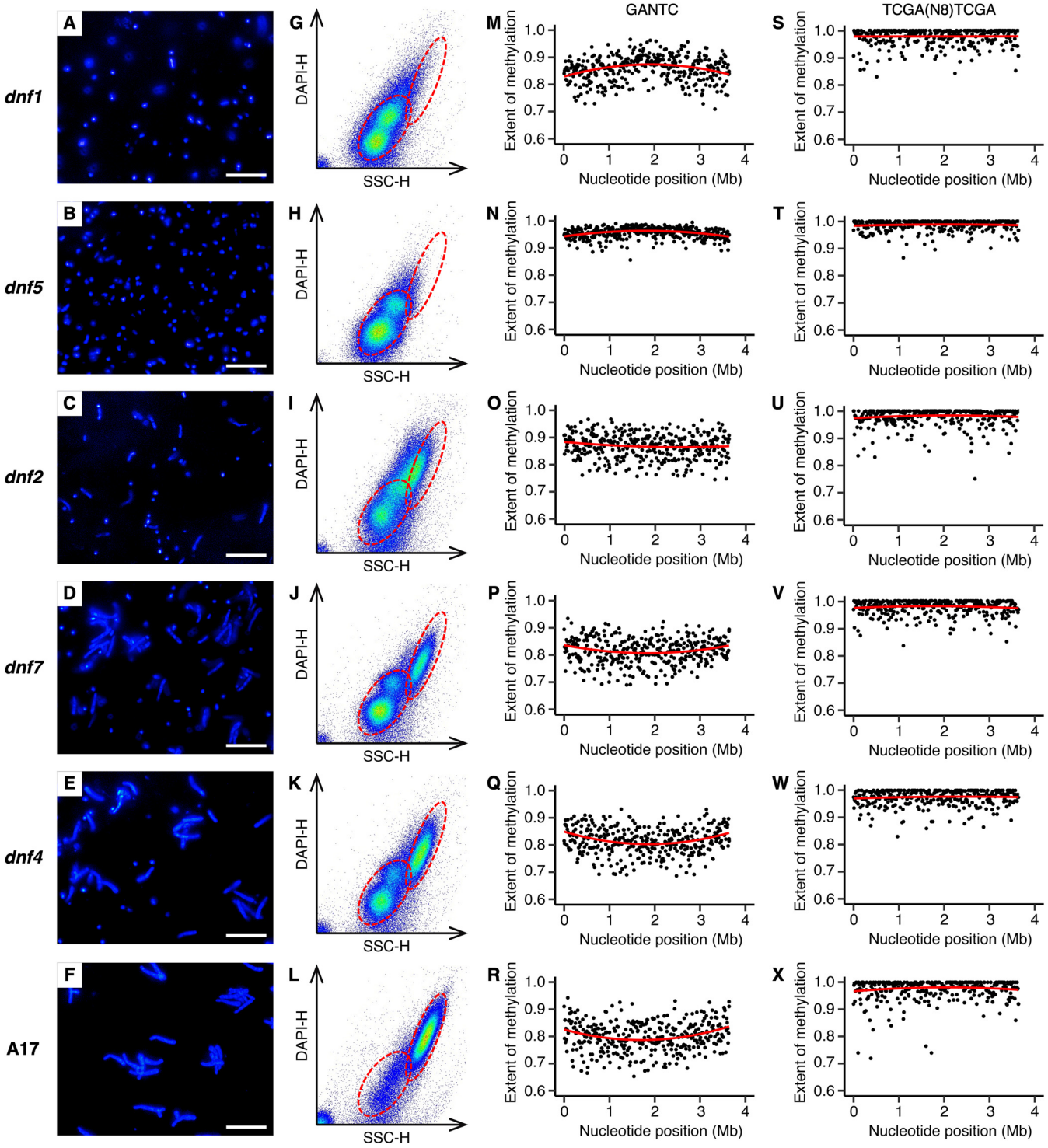


FIG 6 Bacteroid morphology and chromosomal GANTC methylation in *E. meliloti* bacteroids purified from *M. truncatula dnf* mutant nodules. (A to X) Data are shown for *E. meliloti* FSM-MA bacteroids purified from (A, G, M, and S) *M. truncatula dnf1* mutant nodules, (B, H, N, and T) *M. truncatula dnf5* mutant nodules, (C, I, O, and U) *M. truncatula dnf2* mutant nodules, (D, J, P, V) *M. truncatula dnf7* mutant nodules, (E, K, Q, and W) *M. truncatula dnf4* mutant nodules, and (F, L, R, and X) *M. truncatula* A17 wild-type nodules. (A to F) Micrographs of *E. meliloti* FSM-MA bacteroids stained with the DNA binding dye DAPI. The scale bar represents 30 μm. (G to L) Pseudocolored scatterplots displaying the cell morphology (*x* axis) and DNA content (*y* axis) of *E. meliloti* FSM-MA bacteroids, as determined based on flow cytometry analysis of DAPI-stained cells. The red dashed ellipses indicate the position of undifferentiated bacteria as in culture (not shown) or in the *dnf1* mutant nodules (lower-left ellipse) or fully mature bacteroids as in the A17 wild-type nodules (top-right ellipse). (M to X) The extent of methylation of (M to R) GANTC or (S to X) TCGA(N_8)TCGA motifs across the *E. meliloti* FSM-MA chromosome, shown using a 10-kb sliding window. The red lines are polynomial regression lines calculated in R using the "rlm" method and the formula "y~poly(x,2)." Data for pSymB and pSymA are shown in Fig. S20 and S21 at doi.org/10.6084/m9.figshare.16556205.

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

TABLE 2 Relative sequencing depth of each *E. meliloti* replicon

Strain	Condition	Relative mean sequencing depth[a]		
		Chromosome	pSymA	pSymB
Rm2011	Mid-exponential phase	1.00 ± 0.00	1.00 ± 0.03	0.98 ± 0.03
Rm2011	Stationary phase	1.00 ± 0.00	1.06 ± 0.02	1.02 ± 0.03
Rm2011	*M. sativa* bacteroids	1.00 ± 0.00	0.64 ± 0.02	0.62 ± 0.02
FSM-MA	Mid-exponential phase	1.00 ± 0.00	1.02 ± 0.03	0.87 ± 0.05
FSM-MA	Stationary phase	1.00 ± 0.00	1.10 ± 0.00	1.01 ± 0.03
FSM-MA	*M. sativa* bacteroids	1.00 ± 0.00	0.79 ± 0.04	0.71 ± 0.02
FSM-MA	*M. truncatula* bacteroids	1.00 ± 0.00	0.66 ± 0.02	0.59 ± 0.01
FSM-MA	*dnf1* bacterial cells	1.00	0.98	0.90
FSM-MA	*dnf5* bacterial dells	1.00	0.96	0.94
FSM-MA	*dnf2* bacterial dells	1.00	0.81	0.73
FSM-MA	*dnf7* bacterial dells	1.00	0.82	0.71
FSM-MA	*dnf4* bacterial dells	1.00	0.80	0.73
FSM-MA	A17 bacteroids	1.00	0.76	0.67

[a]Sequencing depth is presented relative to the sequencing depth of the chromosome in the same sample. Values are the means of triplicate samples ± standard deviation, except for the lower half of the table, for which numbers are based on a single replicate.

The nodule bacteria purified from *M. truncatula dnf2* mutant nodules were a mix of undifferentiated and partially differentiated bacteroids, which were polyploid to an extent similar to bacteroids purified from wild-type A17 nodules (Fig. 6I compared to Fig. 6L); however, their cell size was much smaller (Fig. 6C compared to Fig. 6F). This was similar to differentiating bacteroids purified from *M. truncatula* and *M. sativa* zone II nodule sections, many of which had high ploidy without a corresponding increase in cell size (Fig. S11 and S13 at doi.org/10.6084/m9.figshare.16556205). The GANTC methylation pattern of bacteroids from *dnf2* nodules (Fig. 6O) was also similar to that of zone II nodule sections. There was a consistently high extent of GANTC methylation across the chromosome averaging 0.870, which was less than that of bacteroids purified from *dnf5* nodules (0.956) but higher than that of bacteroids purified from wild-type A17 nodules (0.804) (Table S5 at doi.org/10.6084/m9.figshare.16556205), and without the smiling pattern. The nodule bacteria purified from *M. truncatula dnf7* and *dnf4* nodules also contained a mix of undifferentiated bacteria and fully differentiated bacteroids (Fig. 6D, E, J, and K), with the number of undifferentiated bacteria greater in *dnf7* nodules compared to *dnf4* nodules. The GANTC methylation pattern of bacteroids purified from *dnf7* and *dnf4* nodules was similar to that of bacteroids purified from A17 nodules (Fig. 6P and Q). Overall, we interpret the data from bacteroids purified from section nodules and *M. truncatula dnf* mutant nodules as suggesting that CcrM is dysregulated during terminal bacteroid differentiation and that CcrM is constitutively active during endoreduplication.

Chromosome, pSymB, and pSymA sequencing depth are unequal in *E. meliloti* bacteroids. We noticed that in each bacteroid sample, the average extent of GANTC methylation for the chromosomes of the two strains was lower (by 0.04 to 0.13) than that of pSymA or pSymB, and unlike the chromosome, the extent of GANTC methylation was relatively constant across pSymA and pSymB (Fig. S14 to S17 compared to Fig. 4 and 5). These results suggest that, unlike in free-living cells, replication of the three replicons is not well coordinated during terminal differentiation. In agreement with this hypothesis, the mean sequencing depth across pSymA and pSymB was on average ~33% lower than that of the chromosome in all replicates of the *E. meliloti* whole-nodule bacteroid samples (Table 2). Similarly, the mean sequencing depth across pSymA and pSymB was on average ~23% lower than that of the chromosome for the polyploid bacterial cell populations purified from *M. truncatula dnf2*, *dnf7*, and *dnf4* mutant nodules, but not for the haploid/diploid bacterial cell populations purified from *M. truncatula dnf1* and *dnf5* mutant nodules (Table 2). Assuming sequencing depth is correlated with copy number, this observation suggests that *E. meliloti* bacteroids carry approximately two copies each of pSymA and pSymB per three copies of the chromosome.

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

DISCUSSION

In this study, we examined the genome-wide DNA methylation patterns in the free-living cells of four *Ensifer* strains, and in bacteroids of two *E. meliloti* strains, and detected a total of six methylated motifs. We were able to predict cognate MTases for most of the motifs based on genome annotations, the exception being the WNCCGATG motif of *E. adhaerens* OV14. The CGCA(N_5)GTG motif of *E. meliloti* Rm2011 is presumably methylated by Smc02296 (HsdM), a predicted m6A MTase belonging to the HsdRSM type I R-M system that is known to be functional and reduce transformation efficiency (58, 59). The RCGCCTC motif of *E. meliloti* Rm2011 is possibly methylated by Smc03763, a predicted cytosine MTase located upstream of the gene *vsr*, which putatively encodes a very short patch repair protein. However, as MTases encoded upstream of *vsr*-like genes generally create m5C modifications—not m4C modification as in the RCGCCTC motif—it is possible that methylation of RCGCCTC may instead occur by an unidentified MTase (60). Neither of these proteins are found in the other three strains examined here. The motifs TCGA(N_8)TCGA of *E. meliloti* FSM-MA and CAGA(N_7)GTTG of *E. fredii* NGR234 are likely methylated by SMB554_16155 and NGR_c01340, respectively, which are 88% identical at the amino acid level and appear to function as part of R-M systems. Homologs of these two proteins are not found in the other two strains.

Except for GANTC, each methylated motif was detected as methylated only in a single strain. Moreover, MTases, apart from CcrM, are not conserved among *E. meliloti* strains. The lack of conservation suggests that most DNA MTases do not have a major regulatory role in the genus *Ensifer*, aside from CcrM and its role in cell cycle regulation. Supporting this conclusion, no motif was enriched in the promoter regions of symbiosis, carbon source, or cell cycle-regulated genes, and we did not detect any motifs that were methylated specifically in bacteroids. However, we cannot rule out that one or more methylated motifs may influence specific gene expression during free-living growth, differentiation, or N_2-fixation through extended motifs or proximity to other promoter elements, similar to the interplay between CcrM and GcrA during cell cycle regulation in *C. crescentus* (20, 46).

As previously published studies have provided evidence for a cell cycle transition occurring during terminal bacteroid differentiation (26), we were particularly interested in CcrM, a cell cycle-regulated MTase that is broadly conserved in the *Alphaproteobacteria*, and its cognate DNA motif, GANTC. By identifying GANTC sites in the promoter regions of a previously determined set of 462 cell cycle-regulated genes (45), we defined a candidate CcrM regulon in *E. meliloti* consisting of 111 genes. However, further studies are required to better delineate the CcrM regulon in *E. meliloti*, as individual GANTC sites are not diagnostic of a specific gene being regulated by CcrM; GANTC sites were found in the promoter regions of 904 transcripts that did not display cell cycle regulation, and the promoter regions of cell cycle-regulated genes were not enriched in GANTC sites relative to the whole *E. meliloti* genome. Studies in *C. crescentus* suggest that the impact of the fully or hemi-methylated status of GANTC sites on gene expression is mediated, at least in part, through modulating the activity of the transcriptional regulator GcrA (20, 61). However, not all promoter sites containing a GANTC motif are regulated by GcrA in *C. crescentus*, with the relationship dependent on an extended YGAKTCK motif and the precise position of this motif relative to other promoter elements (61). Likely, CcrM-mediated gene regulation in the genus *Ensifer* is also dependent upon additional sequence elements beyond the GANTC motif. Nevertheless, as a recent study in the alphaproteobacterium *Brevundimonas subvibrioides* suggested (46), the CcrM and GcrA regulons as well as the GcrA-binding motif can show very little overlap between species, and therefore they will have to be defined experimentally in *E. meliloti*.

Consistent with past observations (18), GANTC sites were underrepresented in the genomes of 157 *Ensifer* strains, particularly within coding regions. More surprising, however, was the strong difference in the frequency of GANTC sites between the previously defined (47) symbiotic and nonsymbiotic clades in the genus *Ensifer*, with the frequency of GANTC sites being ~60% higher in the symbiotic clade. Although further work is required to understand the biological significance of the greater frequency of GANTC sites in the symbiotic clade, it is tempting to speculate that it is associated with legume symbiosis.

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

Our data are consistent with CcrM activity differing during terminal bacteroid differentiation compared to free-living cells. The overall moderate to high rates of GANTC methylation in all *E. meliloti* bacteroid samples, coupled with the lack of a chromosome-wide pattern in the *E. meliloti* FSM-MA zone II sample, leads us to hypothesize that CcrM remains constitutively active throughout most of terminal differentiation. This hypothesis is supported by the results for nodule bacteria purified from *M. truncatula dnf* mutant nodules, which showed that differentiation is preceded by full GANTC methylation and that GANTC methylation remains high (but moderately lower) during endoreduplication, followed by another moderate drop in GANTC methylation in late stages of differentiation. Considering that overexpression of CcrM can give rise to bacteroid-like morphology in free-living cells (14), we hypothesize that constitutive CcrM MTase activity is one (of potentially multiple) factor(s) driving polyploidization of bacteroids (Fig. 2). However, further studies monitoring CcrM abundance and artificially manipulating *ccrM* expression throughout bacteroid differentiation are required to conclusively determine if CcrM is constitutively active during terminal differentiation and the importance of this activity to the promotion of endoreduplication.

CcrM activity is confined to a short window in the cell cycle since the *ccrM* gene is expressed in the late phase of genome replication (45) and, at least in *C. crescentus*, the CcrM protein is degraded by the Lon protease prior to cell division (62). Thus, constitutive CcrM MTase activity in differentiating bacteroids could be obtained through an aberrant expression of the gene or, alternatively, a lack of proteolytic degradation of the CcrM protein. In agreement with the latter possibility, Lon protease was identified as a target of the NCR247 peptide (34). NCR247 is one of the ~600 NCR peptides produced by *M. truncatula* that induce bacteroid terminal differentiation, and studies suggest that NCR247 does so at least in part by perturbing expression of genes involved in cell cycle regulation (33). It is tempting to speculate that NCR peptides such as NCR247 inhibit Lon protease activity posttranslationally, thereby stabilizing CcrM and triggering bacteroid differentiation. However, the CcrM MTase does not appear to remain active in fully differentiated bacteroids, with the lower GANTC methylation near the chromosomal *ter* regions suggesting that loss of CcrM MTase activity occurs slightly prior to completion of genome endoreduplication (model provided as Fig. 2). These hypotheses are consistent with *M. truncatula-E. meliloti* nodule zone-specific RNA-sequencing data (24), which showed that *ccrM* expression in the root distal portion of zone II is ~2-fold higher than in the root proximal portion of zone II and ~10-fold higher than in zone III. The ~10-fold difference in *ccrM* expression across nodule zones suggests to us that the level of *ccrM* expression during early stages of bacteroid differentiation is biologically significant, a prerequisite for the constitutive CcrM activity that we hypothesize.

Our analyses also provide insight into the genome replication dynamics of *E. meliloti* during free-living growth and terminal bacteroid differentiation. Notably, flow cytometry data of *E. meliloti* bacteroids purified from zone II nodule sections and *M. truncatula dnf2* nodules suggest that endoreduplication and cell enlargement largely occur subsequently, not concurrently. Genome replication might be a much faster process than cell growth or, alternatively, endoreduplication might be required to drive cell enlargement. Moreover, our data are consistent with a loss of coordination of replication of the three replicons during terminal bacteroid differentiation, leading to unequal copy numbers, with two copies of pSymA and pSymB per three copies of the chromosome in bacteroids. This differs from free-living cells, where the copy number of the three replicons was approximately equal based on average sequencing depth. We hypothesize that this is a consequence of pSymA and pSymB encoding their own replication proteins and having independent regulation of replication initiation and copy number from that of the chromosome (63, 64). The relative change in replicon copy number occurs concomitantly with differentiation and polyploidization, as supported by the relative abundance of the replicons differing in nodule bacterial cells that have experienced endoreduplication (i.e., cells retrieved from *M. truncatula dnf2*, *dnf7*, and *dnf4* mutant nodules) but not in bacterial cells that had not yet undergone endoreduplication (i.e., cells retrieved from *M. truncatula dnf1* and *dnf5* mutant nodules). In contrast to our results, a previous comparison of the relative abundance of the three replicons in free-living

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

E. meliloti Rm1021 versus bacteroids did not detect differences using comparative genome hybridization with microarrays (26). We believe that the difference between our present data and the previous analysis is due to the subtlety of the differences and the lower sensitivity of the microarray hybridization method compared to high-throughput sequencing.

We also observed that during free-living exponential growth, the extent of GANTC methylation at the *ori* of pSymA and pSymB is higher than at the *ori* of the chromosome, while the *ter* of the pSymA and pSymB has a slightly lower extent of GANTC methylation than the *ter* of the chromosome. As GANTC methylation occurs at a fixed stage of the cell cycle corresponding to the end of chromosome replication, our observations indicate that pSymA and pSymB replication is initiated later in the cell cycle than initiation of chromosome replication, while their replication terminates slightly before completion of chromosomal replication and the activation of CcrM (model provided as Fig. 2). These results provide additional support for the notion of spatiotemporal regulation of DNA replication and partitioning in the multipartite *E. meliloti* genome as proposed previously (45, 65). Similarly, replication of chromosome II of *Vibrio cholerae* is delayed relative to chromosome I, leading to the replication of these two replicons terminating at approximately the same time (66). Thus, coordinating the timing or replication termination may be a general feature of multipartite genomes.

MATERIALS AND METHODS

Experimental design. The overall experimental design is summarized in Fig. S1 at doi.org/10.6084/m9.figshare.16556205. Genomic DNA was isolated from four wild-type *Ensifer* strains to explore how DNA methylation varies across this genus; to allow direct comparison, the four strains were grown to mid-exponential phase in minimal medium with succinate as a carbon source. To investigate how DNA methylation patterns differ between actively dividing and nondividing cells, genomic DNA was isolated from *E. meliloti* Rm2011 grown to either mid-exponential phase or stationary phase. Genomic DNA was isolated from *E. meliloti* Rm2011 grown to mid-exponential phase with either a glycolytic (sucrose) or gluconeogenic (succinate) carbon source to investigate whether DNA methylation may play a role in regulating carbon metabolism. Furthermore, an *E. meliloti* Rm2011 derivative lacking the pSymA and pSymB replicons (named RmP3496) was studied to gain insight into whether these replicons contribute to DNA methylation patterns in *E. meliloti* Rm2011; this strain was grown with sucrose (instead of succinate), as RmP3496 lacks the succinate transporter.

In addition to the free-living samples, *E. meliloti* bacteroid samples purified from legume nodules were collected to investigate changes in DNA methylation during bacteroid differentiation and nitrogen fixation. To do so, *E. meliloti* Rm2011 and *E. meliloti* FSM-MA bacteroids were isolated from *M. sativa* whole nodules. *E. meliloti* FSM-MA bacteroids were additionally purified from *M. truncatula* whole nodules to examine the impact of the host plant on bacteroid DNA methylation patterns. *E. meliloti* Rm2011 bacteroids were only isolated from *M. sativa* nodules, as unlike FSM-MA, Rm2011 forms a poor symbiosis with *M. truncatula* (48, 49). Moreover, *E. meliloti* Rm2011 and *E. meliloti* FSM-MA bacteroids were isolated from *M. sativa* nodule sections (sectioned at the white-pink border to separate the root distal infection and differentiation zone II [white] from the root proximal nitrogen-fixing zone III [pink]) to facilitate an analysis of how DNA methylation patterns differ between differentiating bacteroids and fully differentiated and nitrogen-fixing bacteroids. This was followed by isolation of *E. meliloti* FSM-MA bacteroids from whole nodules of *M. truncatula* mutant lines (*dnf1*, *dnf2*, *dnf4*, *dnf5*, *dnf7*) to investigate DNA methylation patterns in nodule bacteria blocked as various stages of differentiation.

Bacterial strains and growth conditions. The bacterial strains used in this work are listed in Table S6 at doi.org/10.6084/m9.figshare.16556205. All strains were routinely grown on tryptone-yeast (TY) with 2 μM $CoCl_2$, as it was required for *E. meliloti* RmP3496 (42). The MM9 minimal medium (42) consisted of the following: 40 mM MOPS (morpholinepropanesulfonic acid), 20 mM KOH, 19.2 mM NH_4Cl, 85.6 mM NaCl, 2 mM KH_2PO_4, 1 mM $MgSO_4$, 0.25 mM $CaCl_2$, 1 μg mL^{-1} biotin, 42 nM $CoCl_2$, 38 μM $FeCl_3$, 10 μM thiamine-HCl, and either 10 mM sucrose (MM9-sucrose) or 20 mM disodium succinate (MM9-succinate). Prior to inoculation of plants with *E. meliloti*, the strains were grown in yeast extract-beef broth (YEB) medium (67).

DNA isolation from free-living cells. Overnight cultures of all strains grown in MM9-succinate or MM9-sucrose medium were diluted into 10 mL of the same medium to a starting optical density at 600 nm (OD_{600nm}) of 0.025 (0.05 for RmP3496) and incubated overnight at 30°C with shaking (130 rpm). The next day, cultures were diluted into 40 mL of the same medium in 100-mL flasks to the OD_{600nm} values listed in Table S7 at doi.org/10.6084/m9.figshare.16556205. To obtain mid-exponential-phase samples, cultures were harvested after 15.5 to 16 h of growth at OD_{600nm} values between 0.37 and 0.69 (Table S7 at doi.org/10.6084/m9.figshare.16556205). To obtain stationary-phase samples, cultures were harvested after 24 h of growth at OD_{600nm} values of ~1.4. In all cases, cultures were streaked on TY plates to check for contamination and then centrifuged (8,200 × *g*, 10 min, 4°C); the full 40 mL was centrifuged for mid-exponential-phase cultures, whereas only 15 mL was centrifuged for stationary-phase cultures. Most of the supernatant was discarded, and the pellet was resuspended in the remaining volume, transferred to a 2-mL tube and centrifuged again (16,200 × *g*, room temperature, 1 min), and the supernatant was discarded. Three biological replicates, each starting from a separate overnight culture, were performed. DNA was isolated using phenol:chloroform extractions

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

and ammonium acetate precipitations as described elsewhere (68), and the DNA pellets (following RNase A treatment) were resuspended in 200 μL of 10 mM Tris-HCl, pH 8.5.

DNA isolation from bacteroids. *M. sativa* cv. Gabès and *M. truncatula* cv. A17 were the wild-type plants used for all experiments. *M. truncatula dnf1*, *dnf2*, *dnf4*, *dnf5*, and *dnf7* mutants (51), derived from the A17 wild type, were used for collection of bacteroids blocked at various stages of differentiation. Seeds were scarified, surface sterilized, and germinated on Kalys agar as described previously (49). First, 50 mL of overnight cultures of *E. meliloti* Rm2011 or FSM-MA, grown in YEB, were centrifuged (4,000 $\times$ *g*, 20 min, room temperature) and resuspended in ~1,200 mL of sterile, distilled water to obtain a cell suspension at an OD_{600nm} of ~0.1. Germinated seedlings were immersed for 1 h in the appropriate rhizobial cell suspension and then planted in a perlite:sand (2:1) mixture. Plants were grown in a greenhouse for five to six weeks, with occasional watering with a 1 g L^{-1} nutrient solution (PlantProd solution [N-P-K, 0-15-40]).

For whole-nodule samples of wild-type plants, pink nodules were collected from 53 to 60 plants per replicate 34- to 35-day postinoculation; in the case of *dnf* mutants (and a matched wild-type A17 sample), nodules were collected from ~105 plants per genotype 23 to 24 days postinoculation. Nodules were collected from the roots and stored in tubes in liquid nitrogen until collection was complete, at which point they were stored at −80°C until use. For sectioned nodule samples, pink nodules were collected from 103 *M. sativa* plants for each of the microsymbionts 35 to 40 days postinoculation. Nodules were manually sectioned at the white-pink border. Nodule sections were stored in tubes over dry ice or liquid nitrogen until collection was complete, at which point they were stored at −80°C until use. Average plant shoot dry weights for all samples are listed in Table S8 at doi.org/10.6084/m9.figshare.16556205. Bacteroids were isolated from the nodule samples using Percoll gradient centrifugation as described elsewhere (26). The recovered bacteroids were resuspended in 50 to 100 μL of bacteroid extraction buffer (BEB; 125 mM KCl, 50 mM Na-succinate, 50 mM TES [*N*-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid], pH 7.0) and either used immediately for microscopy, flow cytometry, and DNA isolation or stored at −80°C until use.

Nucleic acids were initially purified from most bacteroid samples using the Epicentre MasterPure complete DNA and RNA purification kit, following the protocol for DNA isolation from cell samples; the exceptions were bacteroids purified from *dnf* mutant nodules (and the matched wild-type A17 sample), for which nucleic acids were isolated by using phenol:chloroform extractions followed by ammonium acetate DNA precipitations as described elsewhere (68). For sectioned-nodule samples, pure DNA was isolated by using the manufacturer's protocol for the complete removal of RNA. For whole-nodule samples, the isolated DNA was further purified by treating the nucleic acid samples with RNase A, after which pure DNA was isolated by using phenol:chloroform extractions followed by ammonium acetate DNA precipitations or, alternatively, using the MasterPure DNA cleanup protocol for the DNA from *dnf* mutant nodules and the matched wild-type A17 sample. In all cases, the final DNA pellets were resuspended in 200 μL of 10 mM Tris-HCl, pH 8.5. Three biological replicates were performed for bacteroids isolated from most whole nodules, whereas only one replicate was performed for bacteroids isolated from sectioned nodules or *dnf* mutant nodules (and the matched wild-type A17 sample) due to low quantities of starting materials.

DNA sequencing, modification detection, and motif analysis. DNA sequencing was performed at the U.S. Department of Energy Joint Genome Institute (JGI) or in-house at the University of Florence (the stationary-phase samples and *dnf* mutant nodules and the matched wild-type A17 sample) using Pacific Biosciences (PacBio) sequencing technology (69). Genomic DNA was sheared to 3 kb using an LS220 (Covaris, Inc., Woburn, MA, USA) or 15 kb (for stationary-phase samples and bacteroids isolated from *dnf* mutant nodules and the matched wild-type A17 sample) using g-TUBEs (Covaris, Inc.). Sheared DNA was treated with exonuclease to remove single-stranded ends and DNA damage repair mix followed by end repair and ligation of barcoded blunt adapters using the SMRTbell template prep kit 2.0 (PacBio, Menlo Park, CA, USA). Libraries were purified with AMPure PB beads (Beckman Coulter, Brea, CA, USA), and three or eight libraries with different barcodes were pooled at equimolar ratios and purified with AMPure PB beads. For most samples, SMRTbell template libraries were prepared using a Sequel binding kit 3.0 (PacBio, Menlo Park, CA, USA) and sequenced on a Sequel instrument using a v3 or v4 sequencing primer, 1 M v3 SMRT cells, and version 3.0 sequencing chemistry with 1 $\times$ 360 or 1 $\times$ 600 sequencing movie run times. The exceptions were the *E. meliloti* Rm2011 zone II and *E. meliloti* FSM-MA zone III bacteroid samples. For these samples, SMRTbell template libraries were prepared using a Sequel II binding kit 2.0 (PacBio) and then sequenced on a Sequel II instrument using the tbd-sample-dependent sequencing primer, 8 M v1 SMRT cells, and version 2.0 sequencing chemistry with 1 $\times$ 900 sequencing movie run times.

DNA modification detection and motif analysis were performed using the SMRT Link software (PacBio, Menlo Park, CA, USA). Briefly, raw reads were filtered using SFilter to remove short reads and reads derived from sequencing adapters. Filtered reads were aligned against the appropriate reference genome (Table S2 at doi.org/10.6084/m9.figshare.16556205) using BLASR (70), and modified sites were then identified through kinetic analysis of the aligned DNA sequence data (71); the number of mapped bases per sample is provided in Table S2 at doi.org/10.6084/m9.figshare.16556205. Modified sites were then grouped into motifs using MotifFinder. These motifs represent the recognition sequences of MTase genes active in the genome (72). Downstream analyses were performed using custom Perl and R scripts.

Flow cytometry. Flow cytometry was performed as described previously (26). Freshly prepared bacteroid samples were diluted in 200 μL of BEB, heat-treated for 10 min in a 70°C water bath, and then stained with the DNA-binding dye diamidino-2-phenylindole (DAPI). Cell size and ploidy level of the bacteroid samples were determined using flow cytometry with a Beckman Coulter CytoFLEX S instrument. Measurements consisted of 50,000 cells. Data analysis was performed using the CytExpert 2.2.0.97 software.

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

Fluorescence microscopy. One μL of each freshly prepared bacteroid sample was mixed with 1 μL of 50 μg mL^{-1} DAPI or with both 1 μL of 50 μg mL^{-1} DAPI and 1 μL of 100 μg mL^{-1} propidium iodide (PI), which are both DNA binding dyes. Samples were visualized at $\times$100 magnification under oil immersion using a Nikon Eclipse 80*i* fluorescence microscope with the NIS-Elements BR 4.00.01 software and a Digital Sight DS-U3 camera.

Phylogenetic analysis. The nucleotide fasta files of representative *Ensifer* species were downloaded from the National Centre for Biotechnology Information (NCBI) genome database. A core gene phylogeny was constructed using a previously prepared pipeline (73) reliant on the use of Roary 3.11.3 (74), Prokka 1.12-beta (75), PRANK 140110 (76), trimAl (77), and RAxML 8.2.9 (78). The phylogeny was visualized with the iTOL webserver (79). Identification of nodulation (*nodABC*) and nitrogenase genes (*nifHDK*) was performed with a published pipeline (73) reliant on the use of HMMER 3.1b2 (80), and the Pfam-A 31.0 (81) and TIGRFAM 15.0 (82) databases.

Data availability. Most sequencing data are available through the JGI Genome Portal (genome.jgi.doe.gov/portal/) under proposal 503835, as well as through the NCBI (for BioSample accession numbers, see Table S2 at doi.org/10.6084/m9.figshare.16556205). The data for stationary-phase cultures and bacteroids isolated from *dnf* mutant nodules are available only through the NCBI (BioProject accession no. PRJNA706182 and PRJNA705832; for BioSample accession numbers, see Table S2 at doi.org/10.6084/m9.figshare.16556205). Raw flow cytometry FCS files are available as Data set S3 at doi.org/10.6084/m9.figshare.16556205. All custom scripts to perform the analyses described in this study are available through GitHub (github.com/diCenzo-Lab/003_2021_Ensifer_DNA_methylation).

ACKNOWLEDGMENTS

We are grateful to E. Mullins (Teagasc, MTA2018233) for *E. adhaerens* OV14 and M. Bourge (I2BC) for help with the flow cytometry experiments.

Most sequencing work was performed at the U.S. Department of Energy (DOE) Joint Genome Institute (JGI), through the CSP New Investigator Program (proposal 503835). The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under contract no. DE-AC02-05CH11231. This work was partially supported by the "MICRO4Legumes" grant (Italian Ministry of Agriculture), by the grant "Dipartimento di Eccellenza 2018–2022" from the Italian Ministry of Education, University and Research (MIUR), and by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada. The present work benefited from the Imagerie-Gif core facility supported by the Agence Nationale de la Recherche (ANR-11-EQPX-0029/Morphoscope, ANR-10-INBS-04/FranceBioImaging, ANR-11-IDEX-0003-02/Saclay Plant Sciences). G.C.D. was supported by an NSERC postdoctoral fellowship and a European Molecular Biology Organization (EMBO) short-term fellowship for part of this work. L.C. was supported by the MICRO4Legumes grant (Italian Ministry of Agriculture). Q.N. was supported by a Ph.D. fellowship from the Paris-Saclay University. J.H.T.C. was supported by an NSERC Undergraduate Summer Research Award. P.M., E.G.B., and B.A. were supported by Saclay Plant Sciences (SPS) and grant ANR-17-CE20-0011 from the Agence Nationale de la Recherche.

REFERENCES

1. Greenberg MVC, Bourc'his D. 2019. The diverse roles of DNA methylation in mammalian development and disease. Nat Rev Mol Cell Biol 20:590–607. https://doi.org/10.1038/s41580-019-0159-6.
2. Zhang H, Lang Z, Zhu J-K. 2018. Dynamics and function of DNA methylation in plants. Nat Rev Mol Cell Biol 19:489–506. https://doi.org/10.1038/s41580-018-0016-z.
3. Tang Y, Gao X-D, Wang Y, Yuan B-F, Feng Y-Q. 2012. Widespread existence of cytosine methylation in yeast DNA measured by gas chromatography/mass spectrometry. Anal Chem 84:7249–7255. https://doi.org/10.1021/ac301727c.
4. Blow MJ, Clark TA, Daum CG, Deutschbauer AM, Fomenkov A, Fries R, Froula J, Kang DD, Malmstrom RR, Morgan RD, Pósfai J, Singh K, Visel A, Wetmore K, Zhao Z, Rubin EM, Korlach J, Pennacchio LA, Roberts RJ. 2016. The epigenomic landscape of prokaryotes. PLoS Genet 12:e1005854. https://doi.org/10.1371/journal.pgen.1005854.
5. Sánchez-Romero MA, Cota I, Casadesús J. 2015. DNA methylation in bacteria: from the methyl group to the methylome. Curr Opin Microbiol 25:9–16. https://doi.org/10.1016/j.mib.2015.03.004.
6. Gopalakrishnan S, Van Emburgh BO, Robertson KD. 2008. DNA methylation in development and human disease. Mutat Res 647:30–38. https://doi.org/10.1016/j.mrfmmm.2008.08.006.
7. Vasu K, Nagaraja V. 2013. Diverse functions of restriction-modification systems in addition to cellular defense. Microbiol Mol Biol Rev 77:53–72. https://doi.org/10.1128/MMBR.00044-12.
8. Atack JM, Tan A, Bakaletz LO, Jennings MP, Seib KL. 2018. Phasevarions of bacterial pathogens: methylomics sheds new light on old enemies. Trends Microbiol 26:715–726. https://doi.org/10.1016/j.tim.2018.01.008.
9. Adhikari S, Curtis PD. 2016. DNA methyltransferases and epigenetic regulation in bacteria. FEMS Microbiol Rev 40:575–591. https://doi.org/10.1093/femsre/fuw023.
10. Campbell JL, Kleckner N. 1990. *E. coli oriC* and the *dnaA* gene promoter are sequestered from *dam* methyltransferase following the passage of the chromosomal replication fork. Cell 62:967–979. https://doi.org/10.1016/0092-8674(90)90271-f.
11. Kang S, Lee H, Han JS, Hwang DS. 1999. Interaction of SeqA and Dam methylase on the hemimethylated origin of *Escherichia coli* chromosomal DNA replication. J Biol Chem 274:11463–11468. https://doi.org/10.1074/jbc.274.17.11463.
12. Lahue R, Au K, Modrich P. 1989. DNA mismatch correction in a defined system. Science 245:160–164. https://doi.org/10.1126/science.2665076.
13. Zweiger G, Marczynski G, Shapiro L. 1994. A *Caulobacter* DNA methyltransferase that functions only in the predivisional cell. J Mol Biol 235:472–485. https://doi.org/10.1006/jmbi.1994.1007.
14. Wright R, Stephens C, Shapiro L. 1997. The CcrM DNA methyltransferase is widespread in the alpha subdivision of proteobacteria, and its essential

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

functions are conserved in *Rhizobium meliloti* and *Caulobacter crescentus*. J Bacteriol 179:5869–5877. https://doi.org/10.1128/jb.179.18.5869-5877.1997.
15. Brilli M, Fondi M, Fani R, Mengoni A, Ferri L, Bazzicalupo M, Biondi EG. 2010. The diversity and evolution of cell cycle regulation in *Alpha-proteobacteria*: a comparative genomic analysis. BMC Syst Biol 4:52. https://doi.org/10.1186/1752-0509-4-52.
16. Kahng LS, Shapiro L. 2001. The CcrM DNA methyltransferase of *Agrobacterium tumefaciens* Is essential, and Its activity is cell cycle regulated. J Bacteriol 183:3065–3075. https://doi.org/10.1128/JB.183.10.3065-3075.2001.
17. Kozdon JB, Melfi MD, Luong K, Clark TA, Boitano M, Wang S, Zhou B, Gonzalez D, Collier J, Turner SW, Korlach J, Shapiro L, McAdams HH. 2013. Global methylation state at base-pair resolution of the *Caulobacter* genome throughout the cell cycle. Proc Natl Acad Sci U S A 110:E4658–E4667. https://doi.org/10.1073/pnas.1319315110.
18. Gonzalez D, Kozdon JB, McAdams HH, Shapiro L, Collier J. 2014. The functions of DNA methylation by CcrM in *Caulobacter crescentus*: a global approach. Nucleic Acids Res 42:3720–3735. https://doi.org/10.1093/nar/gkt1352.
19. Gonzalez D, Collier J. 2013. DNA methylation by CcrM activates the transcription of two genes required for the division of *Caulobacter crescentus*. Mol Microbiol 88:203–218. https://doi.org/10.1111/mmi.12180.
20. Fioravanti A, Fumeaux C, Mohapatra SS, Bompard C, Brilli M, Frandi A, Castric V, Villeret V, Viollier PH, Biondi EG. 2013. DNA binding of the cell cycle transcriptional regulator GcrA depends on N6-adenosine methylation in *Caulobacter crescentus* and other *Alphaproteobacteria*. PLoS Genet 9:e1003541. https://doi.org/10.1371/journal.pgen.1003541.
21. Wang ET, Young JPW. 2019. History of rhizobial taxonomy, p 23–39. *In* Wang ET, Tian CF, Chen WF, Young JPW, Chen WX (ed), Ecology and evolution of rhizobia: principles and applications. Springer, Singapore.
22. Oldroyd GED. 2013. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol 11:252–263. https://doi.org/10.1038/nrmicro2990.
23. Barnett MJ, Toman CJ, Fisher RF, Long SR. 2004. A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. Proc Natl Acad Sci U S A 101:16636–16641. https://doi.org/10.1073/pnas.0407269101.
24. Roux B, Rodde N, Jardinaud M-F, Timmers T, Sauviac L, Cottret L, Carrère S, Sallet E, Courcelle E, Moreau S, Debellé F, Capela D, de Carvalho-Niebel F, Gouzy J, Bruand C, Gamas P. 2014. An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. Plant J 77:817–837. https://doi.org/10.1111/tpj.12442.
25. diCenzo GC, Tesi M, Pfau T, Mengoni A, Fondi M. 2020. Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. Nat Commun 11:2574. https://doi.org/10.1038/s41467-020-16484-2.
26. Mergaert P, Uchiumi T, Alunni B, Evanno G, Cheron A, Catrice O, Mausset A-E, Barloy-Hubler F, Galibert F, Kondorosi A, Kondorosi E. 2006. Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. Proc Natl Acad Sci U S A 103:5230–5235. https://doi.org/10.1073/pnas.0600912103.
27. Czernic P, Gully D, Cartieaux F, Moulin L, Guefrachi I, Patrel D, Pierre O, Fardoux J, Chaintreuil C, Nguyen P, Gressent F, Da Silva C, Poulain J, Wincker P, Rofidal V, Hem S, Barrière Q, Arrighi J-F, Mergaert P, Giraud E. 2015. Convergent evolution of endosymbiont differentiation in dalbergioid and inverted repeat-lacking clade legumes mediated by nodule-specific cysteine-rich peptides. Plant Physiol 169:1254–1265. https://doi.org/10.1104/pp.15.00584.
28. Pini F, Frage B, Ferri L, De Nisco NJ, Mohapatra SS, Taddei L, Fioravanti A, Dewitte F, Galardini M, Brilli M, Villeret V, Bazzicalupo M, Mengoni A, Walker GC, Becker A, Biondi EG. 2013. The DivJ, CbrA and PleC system controls DivK phosphorylation and symbiosis in *Sinorhizobium meliloti*. Mol Microbiol 90:54–71. https://doi.org/10.1111/mmi.12347.
29. Kobayashi H, De Nisco NJ, Chien P, Simmons LA, Walker GC. 2009. *Sinorhizobium meliloti* CpdR1 is critical for co-ordinating cell cycle progression and the symbiotic chronic infection. Mol Microbiol 73:586–600. https://doi.org/10.1111/j.1365-2958.2009.06794.x.
30. Pini F, De Nisco NJ, Ferri L, Penterman J, Fioravanti A, Brilli M, Mengoni A, Bazzicalupo M, Viollier PH, Walker GC, Biondi EG. 2015. Cell cycle control by the master regulator CtrA in *Sinorhizobium meliloti*. PLoS Genet 11:e1005232. https://doi.org/10.1371/journal.pgen.1005232.
31. Gibson KE, Campbell GR, Lloret J, Walker GC. 2006. CbrA is a stationary-phase regulator of cell surface physiology and legume symbiosis in *Sinorhizobium meliloti*. J Bacteriol 188:4508–4521. https://doi.org/10.1128/JB.01923-05.
32. Van de Velde W, Zehirov G, Szatmari A, Debreczeny M, Ishihara H, Kevei Z, Farkas A, Mikulass K, Nagy A, Tiricz H, Satiat-Jeunemaître B, Alunni B, Bourge M, Kucho K, Abe M, Kereszt A, Maróti G, Uchiumi T, Kondorosi E, Mergaert P. 2010. Plant peptides govern terminal differentiation of bacteria in symbiosis. Science 327:1122–1126. https://doi.org/10.1126/science.1184057.
33. Penterman J, Abo RP, De Nisco NJ, Arnold MFF, Longhi R, Zanda M, Walker GC. 2014. Host plant peptides elicit a transcriptional response to control the *Sinorhizobium meliloti* cell cycle during symbiosis. Proc Natl Acad Sci U S A 111:3561–3566. https://doi.org/10.1073/pnas.1400450111.
34. Farkas A, Maróti G, Dürgő H, Györgypál Z, Lima RM, Medzihradszky KF, Kereszt A, Mergaert P, Kondorosi E. 2014. *Medicago truncatula* symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. Proc Natl Acad Sci U S A 111:5183–5188. https://doi.org/10.1073/pnas.1404169111.
35. Satgé C, Moreau S, Sallet E, Lefort G, Auriac M-C, Remblière C, Cottret L, Gallardo K, Noirot C, Jardinaud M-F, Gamas P. 2016. Reprogramming of DNA methylation is critical for nodule development in *Medicago truncatula*. Nat Plants 2:16166. https://doi.org/10.1038/nplants.2016.166.
36. Nagymihály M, Veluchamy A, Györgypál Z, Ariel F, Jégu T, Benhamed M, Szűcs A, Kereszt A, Mergaert P, Kondorosi É. 2017. Ploidy-dependent changes in the epigenome of symbiotic cells correlate with specific patterns of gene expression. Proc Natl Acad Sci U S A 114:4543–4548. https://doi.org/10.1073/pnas.1704211114.
37. Pecrix Y, Staton SE, Sallet E, Lelandais-Brière C, Moreau S, Carrère S, Blein T, Jardinaud M-F, Latrasse D, Zouine M, Zahm M, Kreplak J, Mayjonade B, Satgé C, Perez M, Cauet S, Marande W, Chantry-Darmon C, Lopez-Roques C, Bouchez O, Bérard A, Debellé F, Muños S, Bendahmane A, Bergès H, Niebel A, Buitink J, Frugier F, Benhamed M, Crespi M, Gouzy J, Gamas P. 2018. Whole-genome landscape of *Medicago truncatula* symbiotic genes. Nat Plants 4:1017–1025. https://doi.org/10.1038/s41477-018-0286-7.
38. Davis-Richardson AG, Russell JT, Dias R, McKinlay AJ, Canepa R, Fagen JR, Rusoff KT, Drew JC, Kolaczkowski B, Emerich DW, Triplett EW. 2016. Integrating DNA methylation and gene expression data in the development of the soybean-*Bradyrhizobium* N_2-fixing symbiosis. Front Microbiol 7:518. https://doi.org/10.3389/fmicb.2016.00518.
39. Barrière Q, Guefrachi I, Gully D, Lamouche F, Pierre O, Fardoux J, Chaintreuil C, Alunni B, Timchenko T, Giraud E, Mergaert P. 2017. Integrated roles of BclA and DD-carboxypeptidase 1 in *Bradyrhizobium* differentiation within NCR-producing and NCR-lacking root nodules. Sci Rep 7:9063. https://doi.org/10.1038/s41598-017-08830-0.
40. diCenzo GC, Zamani M, Checcucci A, Fondi M, Griffitts J, Finan TM, Mengoni A. 2019. Multidisciplinary approaches for studying rhizobium-legume symbioses. Can J Microbiol 65:1–33. https://doi.org/10.1139/cjm-2018-0377.
41. diCenzo GC, Muhammed Z, Østerås M, O'Brien SAP, Finan TM. 2017. A key regulator of the glycolytic and gluconeogenic central metabolic pathways in *Sinorhizobium meliloti*. Genetics 207:961–974. https://doi.org/10.1534/genetics.117.300212.
42. diCenzo GC, MacLean AM, Milunovic B, Golding GB, Finan TM. 2014. Examination of prokaryotic multipartite genome evolution through experimental genome reduction. PLoS Genet 10:e1004742. https://doi.org/10.1371/journal.pgen.1004742.
43. Mohapatra SS, Fioravanti A, Biondi EG. 2014. DNA methylation in *Caulobacter* and other *Alphaproteobacteria* during cell cycle progression. Trends Microbiol 22:528–535. https://doi.org/10.1016/j.tim.2014.05.003.
44. Nouri H, Monnier A-F, Fossum-Raunehaug S, Maciag-Dorszynska M, Cabin-Flaman A, Képès F, Wegrzyn G, Szalewska-Palasz A, Norris V, Skarstad K, Janniere L. 2018. Multiple links connect central carbon metabolism to DNA replication initiation and elongation in *Bacillus subtilis*. DNA Res 25:641–653. https://doi.org/10.1093/dnares/dsy031.
45. De Nisco NJ, Abo RP, Wu CM, Penterman J, Walker GC. 2014. Global analysis of cell cycle gene expression of the legume symbiont *Sinorhizobium meliloti*. Proc Natl Acad Sci U S A 111:3217–3224. https://doi.org/10.1073/pnas.1400421111.
46. Adhikari S, Erill I, Curtis PD. 2021. Transcriptional rewiring of the GcrA/CcrM bacterial epigenetic regulatory system in closely related bacteria. PLoS Genet 17:e1009433. https://doi.org/10.1371/journal.pgen.1009433.
47. Fagorzi C, Ilie A, Decorosi F, Cangioli L, Viti C, Mengoni A, diCenzo GC. 2020. Symbiotic and nonsymbiotic members of the genus *Ensifer* (syn. *Sinorhizobium*) are separated into two clades based on comparative genomics and high-throughput phenotyping. Genome Biol Evol 12:2521–2534. https://doi.org/10.1093/gbe/evaa221.
48. Moreau D, Voisin A-S, Salon C, Munier-Jolain N. 2008. The model symbiotic association between *Medicago truncatula* cv. Jemalong and *Rhizobium meliloti* strain 2011 leads to N-stressed plants when symbiotic N2

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

fixation is the main N source for plant growth. J Exp Bot 59:3509–3522. https://doi.org/10.1093/jxb/ern203.
49. Kazmierczak T, Nagymihaly M, Lamouche F, Barrière Q, Guefrachi I, Alunni B, Ouadghiri M, Ibijbijen J, Kondorosi E, Mergaert P, Gruber V. 2017. Specific host-responsive associations between *Medicago truncatula* accessions and *Sinorhizobium strains*. Mol Plant Microbe Interact 30:399–409. https://doi.org/10.1094/MPMI-01-17-0009-R.
50. Vasse J, De Billy F, Camut S, Truchet G. 1990. Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. J Bacteriol 172:4295–4306. https://doi.org/10.1128/jb.172.8.4295-4306.1990.
51. Starker CG, Parra-Colmenares AL, Smith L, Mitra RM, Long SR. 2006. Nitrogen fixation mutants of *Medicago truncatula* fail to support plant and bacterial symbiotic gene expression. Plant Physiol 140:671–680. https://doi.org/10.1104/pp.105.072132.
52. Wang D, Griffitts J, Starker C, Fedorova E, Limpens E, Ivanov S, Bisseling T, Long S. 2010. A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. Science 327:1126–1129. https://doi.org/10.1126/science.1184096.
53. Bourcy M, Brocard L, Pislariu CI, Cosson V, Mergaert P, Tadege M, Mysore KS, Udvardi MK, Gourion B, Ratet P. 2013. *Medicago truncatula* DNF2 is a PI-PLC-XD-containing protein required for bacteroid persistence and prevention of nodule early senescence and defense-like reactions. New Phytol 197:1250–1261. https://doi.org/10.1111/nph.12091.
54. Kim M, Chen Y, Xi J, Waters C, Chen R, Wang D. 2015. An antimicrobial peptide essential for bacterial survival in the nitrogen-fixing symbiosis. Proc Natl Acad Sci U S A 112:15238–15243. https://doi.org/10.1073/pnas.1500123112.
55. Domonkos A, Horvath B, Marsh JF, Halasz G, Ayaydin F, Oldroyd GED, Kalo P. 2013. The identification of novel loci required for appropriate nodule development in *Medicago truncatula*. BMC Plant Biol 13:157. https://doi.org/10.1186/1471-2229-13-157.
56. Horváth B, Domonkos Á, Kereszt A, Szűcs A, Ábrahám E, Ayaydin F, Bóka K, Chen Y, Chen R, Murray JD, Udvardi MK, Kondorosi É, Kaló P. 2015. Loss of the nodule-specific cysteine rich peptide, NCR169, abolishes symbiotic nitrogen fixation in the *Medicago truncatula dnf7* mutant. Proc Natl Acad Sci U S A 112:15232–15237. https://doi.org/10.1073/pnas.1500777112.
57. Lang C, Long SR. 2015. Transcriptomic analysis of *Sinorhizobium meliloti* and *Medicago truncatula* symbiosis using nitrogen fixation–deficient nodules. Mol Plant Microbe Interact 28:856–868. https://doi.org/10.1094/MPMI-12-14-0407-R.
58. Ferri L, Gori A, Biondi EG, Mengoni A, Bazzicalupo M. 2010. Plasmid electroporation of *Sinorhizobium* strains: the role of the restriction gene *hsdR* in type strain Rm1021. Plasmid 63:128–135. https://doi.org/10.1016/j.plasmid.2010.01.001.
59. Brumwell SL, MacLeod MR, Huang T, Cochrane RR, Meaney RS, Zamani M, Matysiakiewicz O, Dan KN, Janakirama P, Edgell DR, Charles TC, Finan TM, Karas BJ. 2019. Designer *Sinorhizobium meliloti* strains and multi-functional vectors enable direct inter-kingdom DNA transfer. PLoS One 14:e0206781. https://doi.org/10.1371/journal.pone.0206781.
60. Bhagwat AS, Lieb M. 2002. Cooperation and competition in mismatch repair: very short-patch repair and methyl-directed mismatch repair in *Escherichia coli*. Mol Microbiol 44:1421–1428. https://doi.org/10.1046/j.1365-2958.2002.02989.x.
61. Haakonsen DL, Yuan AH, Laub MT. 2015. The bacterial cell cycle regulator GcrA is a σ^{70} cofactor that drives gene expression from a subset of methylated promoters. Genes Dev 29:2272–2286. https://doi.org/10.1101/gad.270660.115.
62. Wright R, Stephens C, Zweiger G, Shapiro L, Alley MR. 1996. *Caulobacter* Lon protease has a critical role in cell-cycle control of DNA methylation. Genes Dev 10:1532–1542. https://doi.org/10.1101/gad.10.12.1532.
63. Cevallos MA, Cervantes-Rivera R, Gutiérrez-Ríos RM. 2008. The *repABC* plasmid family. Plasmid 60:19–37. https://doi.org/10.1016/j.plasmid.2008.03.001.
64. Fournes F, Val M-E, Skovgaard O, Mazel D. 2018. Replicate once per cell cycle: replication control of secondary chromosomes. Front Microbiol 9:1833. https://doi.org/10.3389/fmicb.2018.01833.
65. Frage B, Döhlemann J, Robledo M, Lucena D, Sobetzko P, Graumann PL, Becker A. 2016. Spatiotemporal choreography of chromosome and megaplasmids in the *Sinorhizobium meliloti* cell cycle. Mol Microbiol 100:808–823. https://doi.org/10.1111/mmi.13351.
66. Rasmussen T, Jensen RB, Skovgaard O. 2007. The two chromosomes of *Vibrio cholerae* are initiated at different time points in the cell cycle. EMBO J 26:3124–3131. https://doi.org/10.1038/sj.emboj.7601747.
67. Krall L, Wiedemann U, Unsin G, Weiss S, Domke N, Baron C. 2002. Detergent extraction identifies different VirB protein subassemblies of the type IV secretion machinery in the membranes of *Agrobacterium tumefaciens*. Proc Natl Acad Sci U S A 99:11405–11410. https://doi.org/10.1073/pnas.172390699.
68. Cowie A, Cheng J, Sibley CD, Fong Y, Zaheer R, Patten CL, Morton RM, Golding GB, Finan TM. 2006. An integrated approach to functional genomics: construction of a novel reporter gene fusion library for *Sinorhizobium meliloti*. Appl Environ Microbiol 72:7156–7167. https://doi.org/10.1128/AEM.01397-06.
69. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, deWinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, et al. 2009. Real-time DNA sequencing from single polymerase molecules. Science 323:133–138. https://doi.org/10.1126/science.1162986.
70. Chaisson MJ, Tesler G. 2012. Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. BMC Bioinform 13:238. https://doi.org/10.1186/1471-2105-13-238.
71. Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, Korlach J, Turner SW. 2010. Direct detection of DNA methylation during single-molecule, real-time sequencing. Nat Methods 7:461–465. https://doi.org/10.1038/nmeth.1459.
72. Clark TA, Murray IA, Morgan RD, Kislyuk AO, Spittle KE, Boitano M, Fomenkov A, Roberts RJ, Korlach J. 2012. Characterization of DNA methyltransferase specificities using single-molecule, real-time DNA sequencing. Nucleic Acids Res 40:e29. https://doi.org/10.1093/nar/gkr1146.
73. diCenzo GC, Debiec K, Krzysztoforski J, Uhrynowski W, Mengoni A, Fagorzi C, Gorecki A, Dziewit L, Bajda T, Rzepa G, Drewniak L. 2018. Genomic and biotechnological characterization of the heavy-metal resistant, arsenic-oxidizing bacterium *Ensifer* sp. M14. Genes 9:379. https://doi.org/10.3390/genes9080379.
74. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693. https://doi.org/10.1093/bioinformatics/btv421.
75. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
76. Löytynoja A. 2014. Phylogeny-aware alignment with PRANK, p 155–170. *In* Russell D (ed), Methods in molecular biology (methods and protocols). Humana Press, Totowa, NJ.
77. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973. https://doi.org/10.1093/bioinformatics/btp348.
78. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.
79. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 44:W242–W245. https://doi.org/10.1093/nar/gkw290.
80. Eddy SR. 2009. A new generation of homology search tools based on probabilistic inference. Genome Inform 23:205–211.
81. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 44:D279–D285. https://doi.org/10.1093/nar/gkv1344.
82. Haft DH, Selengut JD, Richter RA, Harkins D, Basu MK, Beck E. 2013. TIGRFAMs and genome properties in 2013. Nucleic Acids Res 41:D387–D395. https://doi.org/10.1093/nar/gks1234.

Downloaded from https://journals.asm.org/journal/msystems on 31 January 2024 by 2001:760:2c05:159:b0d4:6a05:b380:a3c.

International Journal of
Molecular Sciences

MDPI

Article

Crossing Bacterial Genomic Features and Methylation Patterns with MeStudio: An Epigenomic Analysis Tool

Christopher Riccardi †, Iacopo Passeri †, Lisa Cangioli, Camilla Fagorzi, Marco Fondi * and Alessio Mengoni *

Department of Biology, University of Florence, Via Madonna del Piano 6, 50019 Sesto Fiorentino, Italy
* Correspondence: marco.fondi@unifi.it (M.F.); alessio.mengoni@unifi.it (A.M.)
† These authors contributed equally to this work.

Abstract: DNA methylation is one of the most observed epigenetic modifications. It is present in eukaryotes and prokaryotes and is related to several biological phenomena, including gene flow and adaptation to environmental conditions. The widespread use of third-generation sequencing technologies allows direct and easy detection of genome-wide methylation profiles, offering increasing opportunities to understand and exploit the epigenomic landscape of individuals and populations. Here, we present a pipeline named MeStudio, with the aim of analyzing and combining genome-wide methylation profiles with genomic features. Outputs report the presence of DNA methylation in coding sequences (CDSs) and noncoding sequences, including both intergenic sequences and sequences upstream of the CDS. We apply this novel tool, showing the usage and performance of MeStudio, on a set of single-molecule real-time sequencing outputs from strains of the bacterial species *Sinorhizobium meliloti*.

Keywords: DNA methylation; mapping; genome sequencing; epigenomics

check for updates

Citation: Riccardi, C.; Passeri, I.; Cangioli, L.; Fagorzi, C.; Fondi, M.; Mengoni, A. Crossing Bacterial Genomic Features and Methylation Patterns with MeStudio: An Epigenomic Analysis Tool. *Int. J. Mol. Sci.* **2023**, *24*, 159. https://doi.org/10.3390/ijms24010159

Academic Editor: Joao Paulo Gomes

Received: 13 October 2022
Revised: 5 December 2022
Accepted: 15 December 2022
Published: 21 December 2022

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Understanding organism adaptation to variable environmental conditions is pivotal for weighting the relevance of natural selection over species and population evolution. Phenotypic plasticity, stress responses, and acclimation contribute significantly to epigenetic mechanisms [1]. Among epigenetic modifications, DNA methylation has been shown to be essential in the control of several biological phenomena in eukaryotes and prokaryotes [2], and, in recent years, the study of variation in epigenetic response aroused the attention of several investigators [3]. Third-generation sequencing technologies, namely, single molecule real-time (SMRT) [4,5] and nanopore ONT [6,7] sequencing, allow rapidly identifying the most commonly methylated bases [8–10]. These methods are improving genome-wide DNA methylation studies, especially in prokaryotes, where the compact size of genomes allows the generation of whole-genome methylomes with relative ease. In prokaryotic microorganisms, DNA methylation plays various roles, which span from control of the cell cycle to protection against phages (e.g., restriction-modification systems) and regulation of gene expression (see, e.g., [11]). Relative to cell cycle control, genome-wide DNA methylation profiles have been shown to vary in ecologically relevant contexts (e.g., bacterial differentiation [12]), as well as for restriction-modification systems with respect to strain or population variation [12].

Consequently, the interest in computational pipelines which can easily profile DNA methylation features in a genome-wide manner (thus allowing a comparison of strains and individuals across multiple conditions) is growing. Several tools have been developed for the analysis of DNA methylation profiles deriving from bisulfite sequencing and microarrays (e.g., [13–17]; for a recent benchmarking, see [18]). Recently, three packages have been released [19–21], which are used to visualize methylation profiles from ONT sequencing

data. A recent tool on GitHub was also developed to specifically analyze DNA methylation profiles on metagenomic data (https://github.com/hoonjeseong/Meta-epigenomics (accessed on 7 September 2022)). However, to the best of our knowledge, no specific pipeline has been developed for extracting DNA methylation information from sequencing data to allow a direct quantification/comparison of the position of methylated sites with respect to genome-derived features, such as coding and noncoding sequences and report outputs, which can be used in population epigenomic analyses. The position of methylated sites with respect to genomic features is of key importance in studies focusing on the role that epigenetic modifications have in gene expression control and phenotypic plasticity.

Here, we report the implementation of a bioinformatic tool, named MeStudio, to explore the methylation profiles and map the methylation patterns to genomic features on a set of genome sequences obtained by SMRT technology of the model symbiotic nitrogen-fixing bacterium *Sinorhizobium meliloti* [22] for which DNA methylation plays a relevant role in cell cycle regulation and differentiation during symbiotic conditions [23]. MeStudio is a pipeline for SMRT sequencing methylation data integration and visualization, combining methylation data with genome sequence and annotation to facilitate the extraction of biological information from DNA methylation profiles. Visual and tabular outputs are produced, which can be further processed to provide biological interpretation and formulate hypotheses on epigenomic profiles.

2. Results and Discussion

2.1. Tool-Wide Comparison

MeStudio provides a novel amount of feature-level information that is not present in other widely used genomic software packages. For instance, Bedtools (https://bedtools.readthedocs.io/en/latest/ (accessed on 17 August 2022)) is a well-known toolset for genomic applications through which it is possible to detect methylation features regarding CpG island, but it is not possible to extract information about CDS, nCDS, tIG, and US regions as it does not provide any figure about methylated motif occurrences. Bioconductor also supplies packages that can be used for methylation analysis such as "*GenomicRanges*" (https://bioconductor.org/packages/release/bioc/html/GenomicRanges.html (accessed on 4 May 2022)) and "*motifmatchr*" (https://bioconductor.org/packages/release/bioc/html/motifmatchr.html (accessed on 6 May 2022)). *GenomicRanges* allows analyzing the genome by dividing it into predefined intervals, but no information about the genomic feature is produced. The package *motifmatchr* finds motifs along the genome, but no gene or protein annotations are included in the output. Moreover, MeStudio simply takes as input for the motifs a text file, which is a more user-friendly format compared to the one required by *motifmatchr*. Table 1 provides a comparison of the features of MeStudio to tools for similar purposes.

Table 1. MeStudio features compared to existing tools.

Tool	Programming Language	Motif Recognition	Motif Matching with Respect to Genomic Features	Graphical Outputs	Reference
MeStudio	Python, C	Yes	Yes	Yes	This study
GenomicRanges	R, C	No	No	Yes	Bioconductor package
motifmatchr	R, C++	Yes	Yes (only providing genomic ranges)	Yes	Bioconductor package
Meta-epigenomics	Python	Yes	No	No	https://github.com/hoonjeseoho/Meta-epigenomics (accessed on 19 June 2022)
Methplotlib	Python, Bash	No	No	Yes	De Coster et al. (2020) [21]
a-slide/pycoMeth	Python, Bash	No	No	Yes	Leger (2020) [20]
NanoMethViz	Python, Bash	No	No	Yes	Su et al. (2021) [19]

2.2. The Sinorhizobium Case Study

In order to show the performance of MeStudio, the genomic sequences of two strains of the model symbiotic nitrogen-fixing bacterium *S. meliloti* were produced and analyzed

together with two additional recently published SMRT data [12] for a total of four genomic sequences of *S. meliloti* strains, 2011, FSM-MA, BE31LL, and BO21CC (Table S1). On the SMRT assembled reads of the genomes of the strains, MeStudio was able to identify a total of 26 motifs (Figure 1). All but six motifs (namely, CTYCCAG, DCTGCAGGS, GCCGGCYD, RAGCWGCTY, RCCAGCC, and RCTGCAGGS) were common to the four strains. The number of retrieved methylated sites ranged from a few units (especially for private motifs, those present in one strain only) to several thousands (such as GANTC, which is a classical motif methylated by the CcrM DNA methylase and involved in cell-cycle regulation [23]. CDS and nCDS showed similar frequencies (Figure 1) (Supplementary Material Table S1), as expected for methylation being present on both DNA strands. Intergenic sequences (tIG) showed the lowest number of methylated sites, while upstream sequences to a gene (US), *bona fide* corresponding to putative promoter regions, reported values generally one order of magnitude higher than tIG, and, in some cases, differences in values between strains ranged around twofold (e.g., CTYCCAG and GCCAGG). Furthermore, differences in the abundance of methylated profiles are evident if we consider the two strains grown until the late exponential phase in minimal medium (i.e., FSM-MA and 2011) and those grown in TY medium (i.e., BE31LL and BO21CC). Lastly, the presence of motifs in one strain only may suggest the occurrence of strain-specific restriction-modification systems, although the small number of methylated sites may also suggest alternative hypotheses (i.e., methylation on some genomic regions only related to regulation of expression at specific loci).

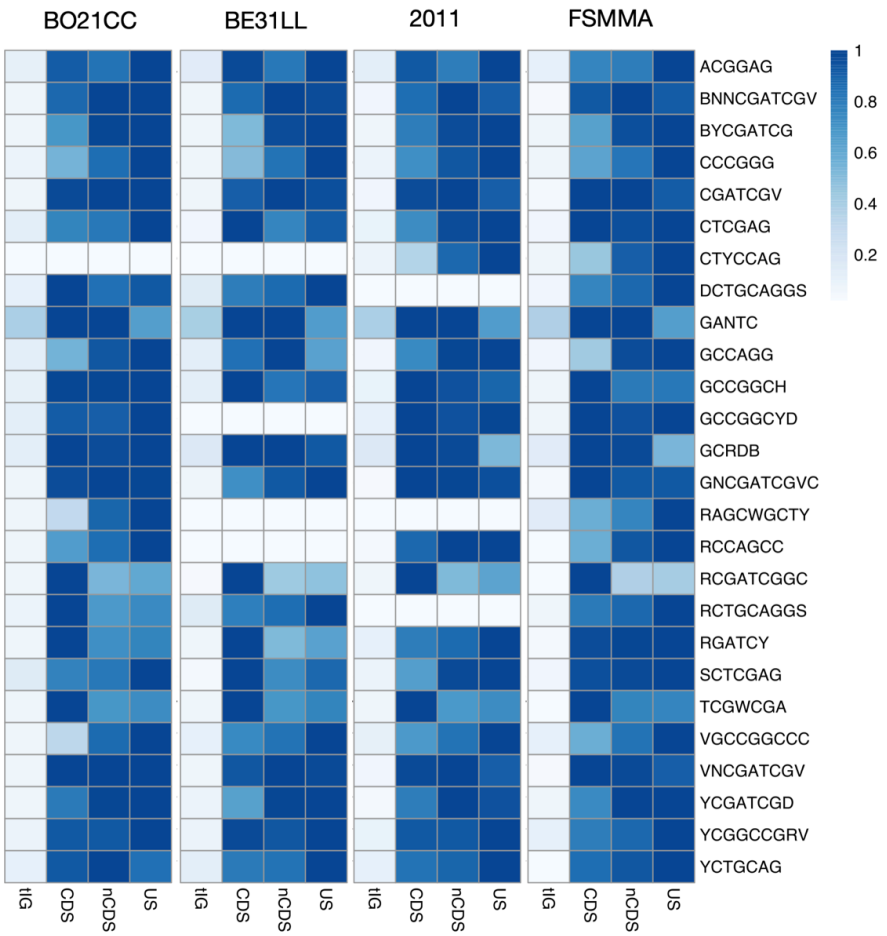


Figure 1. Heatmap representing methylated motif's occurrences scaled as ratio to the max.

In conclusion, we encourage the use of MeStudio to unearth epigenomic data which are interpretable in a comparative genomic framework: the correlation among a methylation position, motifs of interest, and the protein annotation related to a CDS region strengthens the inference between the epigenetic modification and its functional role.

3. Materials and Methods

3.1. Bacterial Strains and Culture Conditions

Strains of *S. meliloti* BE31LL and BO21CC were resuscitated from glycerol stock tubes (codes BM932, BM936) stored at −80 °C in the collection of the Laboratory of Microbial Genetics, Dep. Of Biology, University of Florence, Italy. After re-isolation on TY medium agar plates [24] (tryptone 5 g/L, yeast extract 0.4 g/L, $CaCl_2$ 0.4 g/L, and agar 7.5 g/L), single colonies were inoculated in 5 mL liquid TY medium and grown under constant agitation (125 rpm) at 30 °C.

3.2. DNA Extraction and SMRT Sequencing

DNA was extracted from overnight cultures (OD_{600nm} = 1.5) using PowerSoil DNA Isolation Kit (Qiagen, Hilden, Germany). After quantification by gel electrophoresis and fluorimetric assay (Qubit, Thermo Fisher Scientific, Waltham, MA, USA), we followed the procedure already reported in [25] for fragmenting DNA with g-TUBE (Covaris Inc., Woburn, MA, USA) to an average 15 kbp size and preparing the sequence library using the Pacific Biosciences SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences, Menlo Park, CA, USA). Sequencing was performed on a Sequel apparatus (Pacific Biosciences, Menlo Park, CA, USA) by SMRT technology [21], using Sequel Sequencing Kit 3.0.

3.3. Sequence Analysis and Annotation

The obtained SMRT reads were assembled using the SMRT Link software ver. 8.0.0.80529 (Pacific Biosciences, Menlo Park, CA, USA), producing oriC-oriented assemblies. Annotation was performed using Prokka v1.14.5 [26]. Sequences were deposited in the NCBI database and are available under accession numbers SAMN16976749 and SAMN16976751 (BioProject PRJNA681719). Two additional genomic sequences were analyzed corresponding to *S. meliloti* strains 1021 and FSM-MA, deposited under BioProject PRJNA705832 [12].

3.4. Software Design and Implementation

MeStudio consists of several tools that can be run individually or as part of a pipeline, and it uses a naïve string-matching algorithm to map motif sequences to the reference genome. The required input data consist of only three files: (i) a FASTA file containing the genome sequence, (ii) a genomic annotation file in GFF3 format, and (iii) another GFF3 containing the methylated nucleotide positions. The latter is automatically generated from the output of the SMRTlink software of Pacific Biosciences DNA sequencers. As a result, MeStudio produces several files including: (i) a text file with summarized statistics of the methylation occurrences along the genomic features, (ii) distribution and circular plots, and (iii) BED files containing protein annotation of the genes in which methylated motifs have been found. A complete workflow is provided in Figure 2. Demo files for input and outputs are available at https://github.com/combogenomics/MeStudio (accessed on 4 October 2022).

3.5. Preprocessing

In the first instance, MeStudio performs a quality check via a preprocessing Python script named *ms_replacR*. For a proper analysis, MeStudio needs consistent formatting on the sequence identifiers of the three main input files: the genomic annotation (GFF3 format), sequencer-produced modified base calls (GFF3 format), and the genomic sequence (FASTA file). Since these files may derive from different sources, it is possible for the user to experience differences in the syntax and/or annotation of the sequence identifiers (the "seqid" field); to avoid these possible inconsistencies, *ms_replacR* copies the original files into the output directory, performs a quality check, and corrects the errors, if needed. The most frequent formatting issue we have encountered is the presence of a pipe character and an underscore used interchangeably and inconsistently across files deriving from several sources. To correct for this possible incompatibility issue, by default, the pipe symbol is

replaced by the underscore as a separator. More details are provided in the MeStudio manual on GitHub.

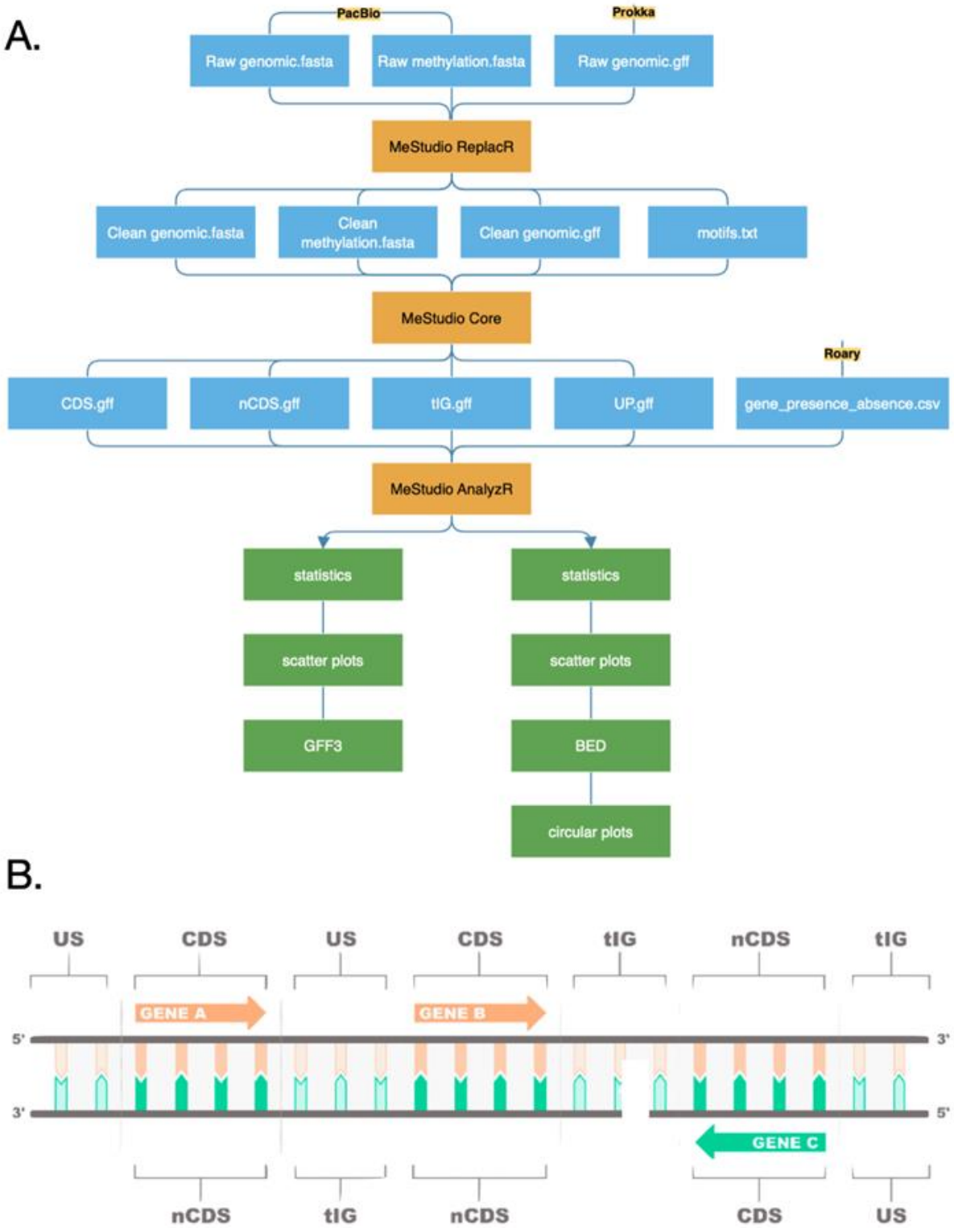


Figure 2. MeStudio overview. (**A**) Workflow. Each blue block represents input files. The orange blocks indicate the scripts. The green boxes indicate output files. (**B**) Graphical representation of the used terminology: CDS, coding sequence; nCDS, coding sequence opposite strand; tIG, intergenic sequence between two genes in opposite directions; US, upstream sequence to a coding sequence (intergenic sequence between two genes having the same orientation). See text for details.

3.6. Core Processing

The processing of the input files is handled by five executables which we refer to as the "MeStudio Core". These components match the nucleotide motifs to the genomic sequence and map them to the corresponding category, which are extracted from the annotation file. Categories are defined as follows: (i) protein-coding genes with an accordant (sense) strand (CDS), (ii) a discordant (antisense) strand (nCDS), (iii) regions that fall between annotated genes (true intergenic, tIG), and (iv) regions upstream of the reading frame of a gene, with an accordant strand (US) (Figure 2B). The CDS feature is defined by the ORF, and the nCDS is its corresponding on the antisense strand. The tIG is defined as the region between two different ORFs on both strands, as reported in Figure 2B. The US region is defined, by default, as the portion of the genome between the end of an ORF and the beginning of the next one; on the other hand, it is possible to set a personalized upstream range via an appropriate flag. The current implementation uses an optimized naive string-matching algorithm to map motif sequences to the reference genome. During the matching stage, each replicon or chromosome is loaded in memory, and both strands are scanned for the presence of the motif sequences, which can obviously hold ambiguous characters. Time complexity in the worst-case scenario is $O(m \times (n - m + 1)) + alpha$, with alpha being an integer proportional to the number and realization of ambiguity characters present in each motif. All the motifs to be searched must be collected by the user and saved in an appropriate newline-delimited text file. The resulting binary files are then processed by another executable that is called for the task at hand. MeStudio Core crosses methylated base positions relative to the reference sequence starting with the previously described features, producing GFF3 files that serve as input for the final analysis stage. This is a computationally expensive part of the pipeline in which multiple nested for loops and calculations are performed. Integrating one motif on a four-contig genome (6,973,268 bp, 23,433 GANTC motif matches) took 0 min 27.116 s on a single AMD Opteron 6380 processor (2.5 GHz).

3.7. Postprocessing

MeStudio implements a postprocessing Python script named *ms_analyzR* which uses MeStudio Core results to produce analytical statistic outputs and return to the user graphical outputs and tables (in the form of BED files), which can be directly parsed using R. In addition, to strengthen the pipeline with comparative genomic analyses, the "gene_presence_abscence.csv" file produced by Roary [27] is needed to define the methylation level and patterns of core and dispensable genome fractions, as well as to annotate the genes-coded proteins. *ms_analyzR* logs the total number of genes found for each category (CDS, nCDS, tIG, and US). Additionally, methylation data are shown, such as (i) the total number of methylated sites, (ii) the total number of methylated genes, (iii) the ID of the most methylated gene (geneID), and (iv) the product of that gene. Integrating data from Roary is functional to characterize the geneID associated with the name of the protein (as annotated by Prokka [26]) as part of the core or dispensable genome. All the information is saved into a log file, together with plots accounting for the distribution of the methylations (Figure 3). To ensure customizability, *ms_analyzR* also includes two optional flags: "—make_chrom" and "—make_bed". The "—make_chrom" flag saves into the previously specified output directory the GFFs at the "chromosome level" rather than the "feature level". Each GFF produced is characterized not by feature (CDS, nCDS, tIG, and US) but by chromosomes (or contigs), maintaining the MeStudio Core-derived contents and layout. The "—make_bed" flag produces a BED file for each feature reporting (i) the *chrom* column, with the name of each chromosome or contig, (ii) the start and (iii) end of the feature, (iv) the name of the geneID found in that interval, (v) the number of methylations found for geneID, and (vi) the protein product of the ID. Information contained in BED files can be readily used to plot the distribution of the methylation density for each feature, making use of the *circlize* R package (https://github.com/jokergoo/circlize (accessed on 18 July 2022)) (Figure 4); an R script for this purpose is already available on our GitHub.

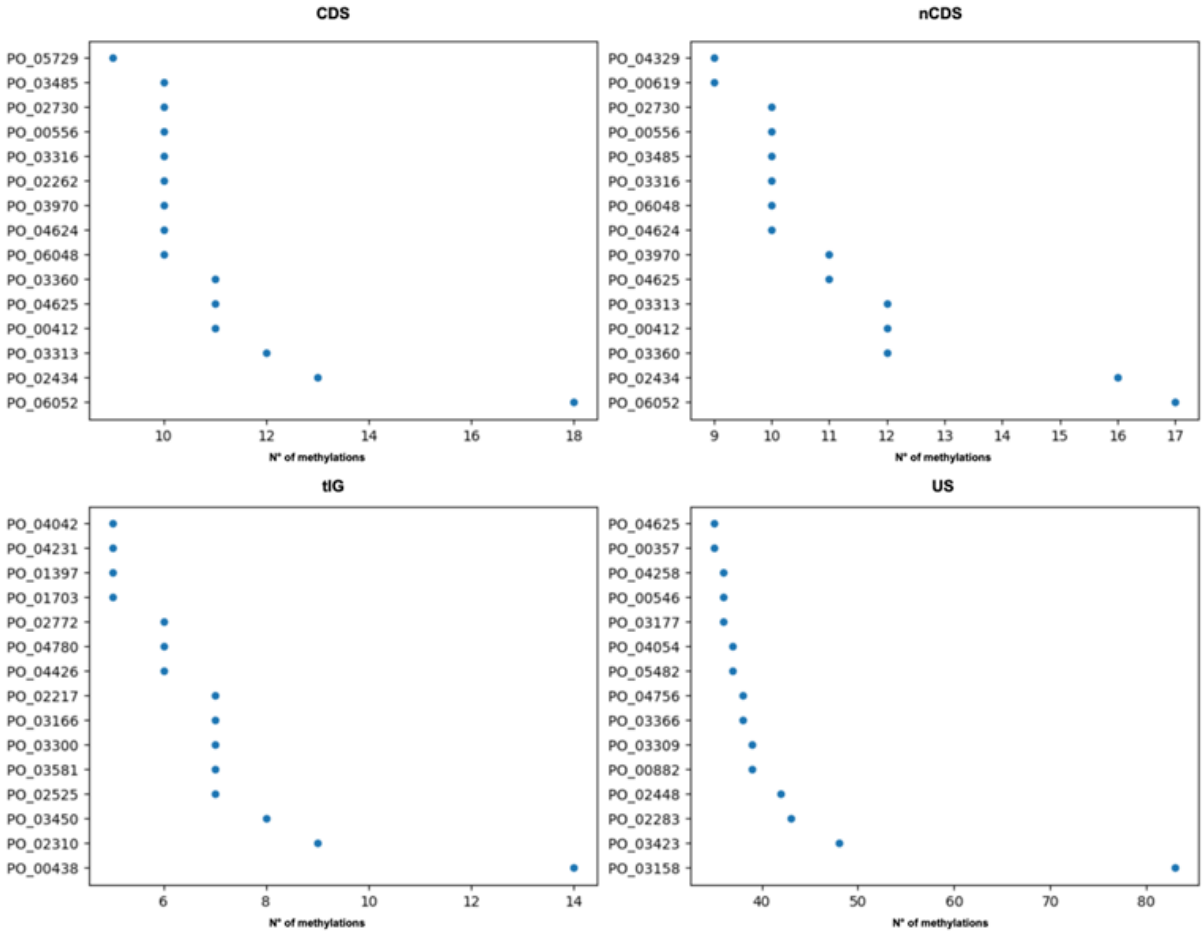


Figure 3. Scatter plots of GANTC motif in *S. meliloti* FSM-MA. The Y-axis reports geneIDs, whereas the X-axis reports the number of methylations found for each geneID. GeneIDs are taken from the annotation (see GitHub repository for the annotation files: https://github.com/combogenomics/MeStudio (accessed on 4 October 2022)). Plots for the different categories of methylated sites (CDS, nCDS, tIG, and US) are reported.

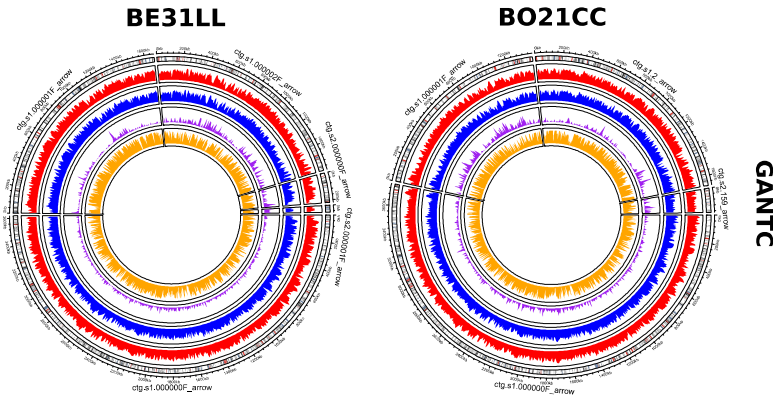


Figure 4. *Cont.*

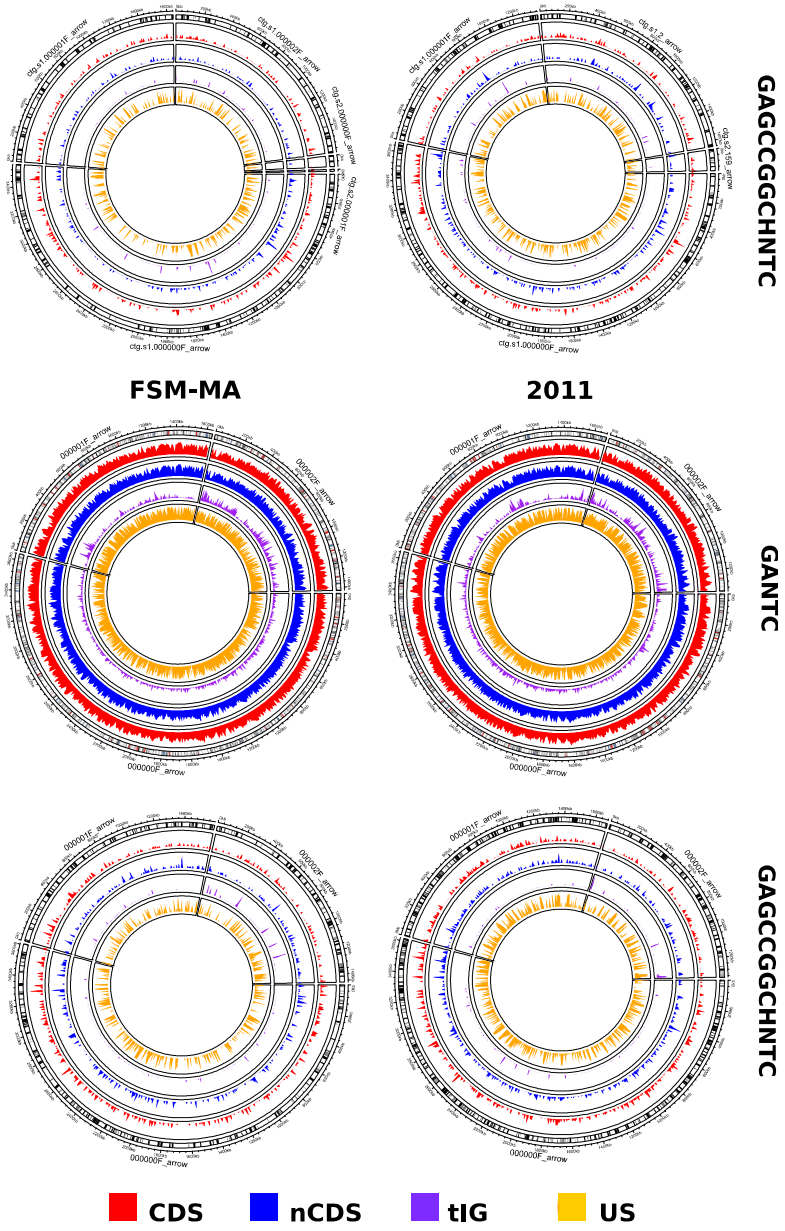


Figure 4. Circular density plots of GANTC and GCCCGGCH motifs in FSM-MA and 1021 strains of *S. meliloti*. The outer circle represents the genome annotation of the contigs of the strain (black lines indicate the position of CDS). Each inner circle represents a different category of methylated sites, CDS (red), nCDS (blue), tIG (purple), and US (yellow). The bars of each plot indicate the values for each category.

4. Conclusions

We reported here the description of a novel software called MeStudio, for the analysis of DNA methylation profiles obtained by single-molecule real-time sequencing. MeStudio has several novel and useful features compared to the few existing tools, as it provides outputs in the form of GFF and BED files which contain information on the position of methylated sites and methylated motifs, the number of methylated sites and profiles for each genomic feature, and graphical outputs, as well as protein annotation. The genomic features analyzed include genic and intergenic regions (comprising putative promoters), allowing the formulation of hypotheses related to the importance of DNA methylation on the regulation of gene expression and on other relevant biological phenomena [28]. In addition to being developed for prokaryotic genomes, MeStudio can handle any kind

of sequence, by simply providing a suitable set of input files (Figure 2A). By providing information on motif occurrence and genomic localization, MeStudio contributes to the basis for comparative analyses of DNA methylation profiles among strains, in terms of evolutionary studies on populations and species, as well as epigenomic modifications during adaptation and development.

Lastly, MeStudio is very user-friendly given its easy installation and its possibility to be run as a pipeline in a single command line call. We developed the scripts in Mac OS and Linux kernel environments, with the possibility of expansion to Windows platforms. Moreover, we plan to make MeStudio affordable to ONT data.

Supplementary Materials: The supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms24010159/s1.

Author Contributions: C.R. and I.P., conceptualization, methodology, software data curation, and writing—original draft preparation; L.C., conceptualization and investigation; C.F., conceptualization and investigation; A.M., conceptualization, and writing—reviewing and editing; M.F., conceptualization, writing—reviewing and editing, validation, and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by MIUR, Programma Nazionale di Ricerche in Antartide 2018, (grant PNRA18_00335, https://www.pnra.aq/) and PRIN (Programmi di Ricerca Scientifica di Rilevante Interesse Nazionale Escaping the ESKAPEs: integrated pipelines for new antibacterial drug, 20208LLXEJ_002) grants to M.F. and by the grant MICRO4Legumes, D.M.n.89267 (Italian Ministry of Agriculture) to A.M. Additionally, L.C. is supported by a PhD fellowship from MICRO4Legumes, while C.F. is supported by a postdoctoral fellowship from the H2020 ERA-NETs SUSFOOD2 and CORE Organic Cofund, under the Joint SUSFOOD2/CORE Organic Call 2019.

Data Availability Statement: The data that support the findings of this study are openly available on GitHub at https://github.com/combogenomics/MeStudio (accessed on 4 October 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Moler, E.R.V.; Abakir, A.; Eleftheriou, M.; Johnson, J.S.; Krutovsky, K.V.; Lewis, L.C.; Ruzov, A.; Whipple, A.V.; Rajora, O.P. Population Epigenomics: Advancing Understanding of Phenotypic Plasticity, Acclimation, Adaptation and Diseases. In *Population Genomics*; Rajora, O.P., Ed.; Population Genomics; Springer International Publishing: Cham, Switzerland, 2018; pp. 179–260, ISBN 978-3-030-04587-6.
2. Jones, P.A. Functions of DNA Methylation: Islands, Start Sites, Gene Bodies and Beyond. *Nat. Rev. Genet.* **2012**, *13*, 484–492. [CrossRef] [PubMed]
3. Chen, P.; Bandoy, D.J.D.; Weimer, B.C. Bacterial Epigenomics: Epigenetics in the Age of Population Genomics. In *The Pangenome: Diversity, Dynamics and Evolution of Genomes*; Tettelin, H., Medini, D., Eds.; Springer: Cham, Switzerland, 2020; ISBN 978-3-030-38280-3.
4. Flusberg, B.A.; Webster, D.R.; Lee, J.H.; Travers, K.J.; Olivares, E.C.; Clark, T.A.; Korlach, J.; Turner, S.W. Direct Detection of DNA Methylation during Single-Molecule, Real-Time Sequencing. *Nat. Methods* **2010**, *7*, 461–465. [CrossRef] [PubMed]
5. Fang, G.; Munera, D.; Friedman, D.I.; Mandlik, A.; Chao, M.C.; Banerjee, O.; Feng, Z.; Losic, B.; Mahajan, M.C.; Jabado, O.J.; et al. Genome-Wide Mapping of Methylated Adenine Residues in Pathogenic Escherichia Coli Using Single-Molecule Real-Time Sequencing. *Nat. Biotechnol.* **2012**, *30*, 1232–1239. [CrossRef] [PubMed]
6. Clarke, J.; Wu, H.-C.; Jayasinghe, L.; Patel, A.; Reid, S.; Bayley, H. Continuous Base Identification for Single-Molecule Nanopore DNA Sequencing. *Nat. Nanotechnol.* **2009**, *4*, 265–270. [CrossRef]
7. Simpson, J.T.; Workman, R.E.; Zuzarte, P.C.; David, M.; Dursi, L.J.; Timp, W. Detecting DNA Cytosine Methylation Using Nanopore Sequencing. *Nat. Methods* **2017**, *14*, 407–410. [CrossRef]
8. Gouil, Q.; Keniry, A. Latest Techniques to Study DNA Methylation. *Essays Biochem.* **2019**, *63*, 639–648. [CrossRef]
9. Sánchez-Romero, M.A.; Casadesús, J. The Bacterial Epigenome. *Nat. Rev. Microbiol.* **2020**, *18*, 7–20. [CrossRef]
10. Rand, A.C.; Jain, M.; Eizenga, J.M.; Musselman-Brown, A.; Olsen, H.E.; Akeson, M.; Paten, B. Mapping DNA Methylation with High-Throughput Nanopore Sequencing. *Nat. Methods* **2017**, *14*, 411–413. [CrossRef]
11. Sánchez-Romero, M.A.; Casadesús, J. Waddington's Landscapes in the Bacterial World. *Front. Microbiol.* **2021**, *12*, 685080. [CrossRef]
12. diCenzo, G.C.; Cangioli, L.; Nicoud, Q.; Cheng, J.H.T.; Blow, M.J.; Shapiro, N.; Woyke, T.; Biondi, E.G.; Alunni, B.; Mengoni, A.; et al. DNA Methylation in Ensifer Species during Free-Living Growth and during Nitrogen-Fixing Symbiosis with *Medicago* Spp. *mSystems* 2022. [CrossRef]

13. Müller, F.; Scherer, M.; Assenov, Y.; Lutsik, P.; Walter, J.; Lengauer, T.; Bock, C. RnBeads 2.0: Comprehensive Analysis of DNA Methylation Data. *Genome Biol.* **2019**, *20*, 55. [CrossRef] [PubMed]
14. Teng, C.-S.; Wu, B.-H.; Yen, M.-R.; Chen, P.-Y. MethGET: Web-Based Bioinformatics Software for Correlating Genome-Wide DNA Methylation and Gene Expression. *BMC Genom.* **2020**, *21*, 375. [CrossRef] [PubMed]
15. Hillary, R.F.; Marioni, R.E. MethylDetectR: A Software for Methylation-Based Health Profiling 2021. Available online: https://wellcomeopenresearch.org/articles/5-283/v2 (accessed on 14 September 2022).
16. Aryee, M.J.; Jaffe, A.E.; Corrada-Bravo, H.; Ladd-Acosta, C.; Feinberg, A.P.; Hansen, K.D.; Irizarry, R.A. Minfi: A Flexible and Comprehensive Bioconductor Package for the Analysis of Infinium DNA Methylation Microarrays. *Bioinformatics* **2014**, *30*, 1363–1369. [CrossRef] [PubMed]
17. Bock, C.; Reither, S.; Mikeska, T.; Paulsen, M.; Walter, J.; Lengauer, T. BiQ Analyzer: Visualization and Quality Control for DNA Methylation Data from Bisulfite Sequencing. *Bioinformatics* **2005**, *21*, 4067–4068. [CrossRef]
18. Nunn, A.; Otto, C.; Stadler, P.F.; Langenberger, D. Comprehensive Benchmarking of Software for Mapping Whole Genome Bisulfite Data: From Read Alignment to DNA Methylation Analysis. *Brief. Bioinform.* **2021**, *22*, bbab021. [CrossRef]
19. Su, S.; Gouil, Q.; Blewitt, M.E.; Cook, D.; Hickey, P.F.; Ritchie, M.E. NanoMethViz: An R/Bioconductor Package for Visualizing Long-Read Methylation Data. *PLOS Comput. Biol.* **2021**, *17*, e1009524. [CrossRef]
20. Leger, A. A-Slide/PycoMeth. 2020. Available online: https://zenodo.org/record/4110144#.Y6LUYxVBxPY (accessed on 3 August 2022). [CrossRef]
21. De Coster, W.; Stovner, E.B.; Strazisar, M. Methplotlib: Analysis of Modified Nucleotides from Nanopore Sequencing. *Bioinformatics* **2020**, *36*, 3236–3238. [CrossRef]
22. Geddes, B.A.; Oresnik, I.J. Physiology, Genetics, and Biochemistry of Carbon Metabolism in the Alphaproteobacterium Sinorhizobium Meliloti. *Can. J. Microbiol.* **2014**, *60*, 491–507. [CrossRef]
23. Fioravanti, A.; Fumeaux, C.; Mohapatra, S.S.; Bompard, C.; Brilli, M.; Frandi, A.; Castric, V.; Villeret, V.; Viollier, P.H.; Biondi, E.G. DNA Binding of the Cell Cycle Transcriptional Regulator GcrA Depends on N6-Adenosine Methylation in Caulobacter Crescentus and Other Alphaproteobacteria. *PLoS Genet.* **2013**, *9*, e1003541. [CrossRef]
24. Beringer, J.E. R Factor Transfer in Rhizobium Leguminosarum. *Microbiology* **1974**, *84*, 188–198. [CrossRef]
25. Bianco, C.; Andreozzi, A.; Romano, S.; Fagorzi, C.; Cangioli, L.; Prieto, P.; Cisse, F.; Niangado, O.; Sidibé, A.; Pianezze, S.; et al. Endophytes from African Rice (*Oryza glaberrima* L.) Efficiently Colonize Asian Rice (*Oryza sativa* L.) Stimulating the Activity of Its Antioxidant Enzymes and Increasing the Content of Nitrogen, Carbon, and Chlorophyll. *Microorganisms* **2021**, *9*, 1714. [CrossRef] [PubMed]
26. Seemann, T. Prokka: Rapid Prokaryotic Genome Annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [CrossRef] [PubMed]
27. Page, A.J.; Cummins, C.A.; Hunt, M.; Wong, V.K.; Reuter, S.; Holden, M.T.G.; Fookes, M.; Falush, D.; Keane, J.A.; Parkhill, J. Roary: Rapid Large-Scale Prokaryote Pan Genome Analysis. *Bioinformatics* **2015**, *31*, 3691–3693. [CrossRef] [PubMed]
28. Mouammine, A.; Collier, J. The Impact of DNA Methylation in Alphaproteobacteria. *Mol. Microbiol.* **2018**, *110*, 1–10. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The pan-epigenome of the symbiotic nitrogen fixing bacterium *Sinorhizobium meliloti* unravels unexpected variability of DNA-methylation profiles in closely related strains

Iacopo Passeri, Lisa Cangioli, Marco Fondi, Alessio Mengoni, Camilla Fagorzi**

[1]*Department of Biology, University of Florence, via Madonna del Piano 6, 50019, Sesto Fiorentino, Italy*

**Corresponding author: alessio.mengoni@unifi.it; camilla.fagorzi@unifi.it*

Abstract

In prokaryotes, DNA methylation has been found to be involved in several mechanisms, such as DNA repair, DNA–protein interactions, gene expression, cell cycle progression and self-DNA recognition (the Restriction-Modification systems). Studies on representatives from the same bacterial species have found that genome-wide DNA methylation patterns can be highly variable and may affect phenotypic variation and gene transfer among closely related strains. However, broader evolutionary studies on such epigenomic variation in bacteria are still scarce.

Here, we addressed this point by performing an epigenomic analysis on 21 strains of the facultative plant symbiotic nitrogen-fixing alphaproteobacterium *Sinorhizobium meliloti*. Strains of these species are characterized by a divided (multipartite) genome structure, including a chromosome, a chromid and a (more recently acquired) megaplasmid. Since these strains display extensive genomic and phenotypic variation, they are good models to test evolutionary hypotheses on the relationships among epigenomic signatures, genome structure evolution and phenotypic switches.

Results showed the presence of a wide pan-epigenome with 16 DNA methylated motifs, including both 4mC and 6mA palindromic and nonpalindromic motifs. While 9 motifs have been found methylated by all strains, the remaining had differential methylation between *S. meliloti* strains, constituting a dispensable epigenome. Differences in frequency of methylation were found among replicons, with the megaplasmid and the additional plasmids displaying several motifs with different methylation frequency with respect to the chromosome and the chromid. Moreover, differences between coding, upstream and intergenic regions, were found, suggesting that DNA methylation at specific motifs may play a role in gene regulation and consequently in phenotypic variability among strains.

Altogether, our data indicate the presence of a large epigenomic diversity in *S. meliloti*, with epigenome signatures differing between replicons, reflecting their timing of evolutionary acquisition in *S. meliloti* genomes and suggesting a role of DNA methylation in the variability of gene expression among strains.

Introduction

Epigenetic information, namely DNA methylation and post-translational modification of histones, has been shown to be pivotal in the control of several biological phenomena in eukaryotes, such as cellular differentiation, development and pathogenesis (Jones 2012). In prokaryotic microorganisms, DNA methylation has been found involved in many molecular mechanisms such as DNA repair, DNA–protein interactions, gene expression, cell cycle progression and self-DNA recognition (the Restriction-Modification systems) (Sánchez-Romero and Casadesús 2020). Novel information gained in recent years on methylated DNA in prokaryotes, especially on Bacteria, is delving into the roles of transcriptional regulation and the formation of phenotypic variants (Oliveira and Fang 2021) (Vasu and Nagaraja 2013). Indeed, N^6-methyladenine (6mA) is known to be involved in many events, spanning from chromosome replication to

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

mismatch repair, conjugal transfer, antibiotic resistance (Sánchez-Romero and Casadesús 2020) bacterial differentiation (diCenzo et al. 2022) and phenotypic plasticity as phase variation (Atack et al. 2018). Besides, stress response and drug transport have been related to the presence of C^5-methyl-cytosine (5mC))(Kahramanoglou et al. 2012; Militello et al. 2014), as N^4-methyl- cytosine (4mC) to pathogenesis (Kumar et al. 2018). Evidence is now supporting the idea that DNA methylation in bacteria may influences phenotypes as virulence and host colonization (Blow et al. 2016; Beaulaurier et al. 2019; Oliveira 2021).

Population studies have shown large differences in R-M systems in clinical isolates of *Staphylococcus epidermidis* and in strains of the plant growth promoting bacterium (PGPB) *Bacillus velezensis* (Reva et al. 2019), where some methylation profiles were found abundant upstream to coding sequences, suggesting they can have regulatory effects on transcription. A recent evolutionary study conducted on *Helicobacter pylori* showed that phase-variable methyltransferase (MTase) genes contribute methylome variation among isolates (Estibariz et al. 2020), indicating that bacterial phenotypic variation and adaptation to variable environmental conditions may rely on epigenetic modification too. Other studies have shown that the genomic levels of 5mC relate to antibiotic resistance (Militello et al. 2014; Yuan et al. 2021). More recently, a pan-epigenome analysis of *Mycoplasma agalactiae* revealed the existence of strain-specific motifs, which allowed the identification of orphan MTases (Dordet-Frisoni et al. 2022), affecting the rate of horizontal gene transfer events among *M. agalactiae* strains. Consequently, it is becoming highly relevant for molecular evolutionary studies in bacteria to investigate the extent of the pan-epigenome and evaluate its role in genome evolution and adaptation.

Facultative symbiotic bacteria are relevant models for the study of environmental adaptation, since they colonize multiple environments coping with dramatically contrasting selective pressure. *Sinorhizobium meliloti* is a facultative plant symbiotic bacterium, belonging to the *Rhizobiaceae* family, able to thrive on soil, colonizing the surrounding of plant roots and entering in symbiosis with host plants (mainly from the Leguminous genera *Medicago*, *Melilotus* and *Trigonella*), where it differentiates into bacteroid, establishing an intracellular symbiosis in specialized root organs called nodules and fixing atmospheric dinitrogen into ammonia (Poole et al. 2018; Cangioli et al. 2021; Ledermann et al. 2021). *S. meliloti* is one of the most studied species for what is concerning symbiotic nitrogen fixation (diCenzo et al. 2019). The genomes of several *S. meliloti* strains have been sequenced and many population-genomic and evolutionary studies have been performed on this species in relation to adaptation and selective pressure for symbiosis (see for instance (Batstone et al. 2020; Batstone et al. 2022; Epstein et al. 2022) and multipartite genome evolution (Galardini, et al. 2013; diCenzo et al. 2016). In fact, the strains of this species often harbour a multipartite (or divided) genome, as for many bacterial species interacting with hosts (diCenzo and Finan 2017; diCenzo, et al. 2019), composed by three main replicons: a chromosome, a chromid and a megaplasmid. The latter contains the essential machinery for symbiosis, while the other two are mainly relevant for growth in soil and in the plant rhizosphere (diCenzo et al. 2016 and 2014). As in the other alphaproteobacterial species (Zweiger et al. 1994), *S. meliloti* cell cycle is controlled by the action of the master regulator CtrA, which in turn expresses its activity in relation to DNA methylation status on a pentanucleotide sequence (GANTC), methylated by the orphan DNA MTase CcrM (Fioravanti et al. 2013; Pini et al. 2015). Recently, we investigated the profile of DNA methylation in two strains of *S. meliloti* along the intracellular differentiation phases in the plant root nodule (diCenzo et al. 2022). In this study, we reported evidence that the CcrM activity on GANTC sequence is dysregulated during symbiosis, further suggesting this dysregulation is a driving factor for intracellular differentiation of *S. meliloti* cells into bacteroids. In this work we noticed the presence of methylated motifs other than GANTC sequence motifs. These additional motifs were found with lower occurrence than GANTC, but interestingly many of them were present in only one of the investigated *S. meliloti* strains, suggesting the presence of a pan-epigenome, composed by a shared set of methylated motifs plus a few unshared or strain-specific motifs. In agreement with the presence of a pan-epigenome, in the same work (diCenzo et al. 2022) we analyzed 24 putative DNA MTases identified in a set of 20 *S. meliloti* genomes and we found that CcrM was the only one in common to all the strains, while many others were in single strains only. These DNA MTases present in the dispensable genome fraction can be the genetic determinants of the unshared methylated motifs (viz. dispensable epigenome). Such motifs could be possibly related to R-M systems and determine the formation of barriers to gene flow. Indeed, in a previous work we have shown for the *S. meliloti*

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1021 strain, that the *hsd*R gene, coding for a putative type-I restriction enzyme, plays a role in the frequency of gene transfer from different donor strains (Ferri et al. 2010), suggesting that the observed differences in DNA methylation profiles and DNA MTases in *S. meliloti* may contribute barriers to gene flow and then to the establishment of subpopulations. However, since evidence for a role of DNA methylation in phenotypic variation in bacteria are increasing, we cannot exclude that both the core and the dispensable epigenome could impact on regulation of gene expression. *S. meliloti* strains, beside displaying relevant differences in genome content because of the typical open pangenome structure (Galardini et al. 2013), also present large phenotypic and transcriptomic variation, both in free-living growth and in symbiotic-related conditions (Biondi et al. 2009; Bellabarba et al. 2021; Fagorzi et al. 2021). We may likely hypothesize that this large variability could also be mirrored by genome-wide DNA methylation variation.

In this work, we examined genome-wide methylation pattern variations in *S. meliloti* strains to: i) assess the presence of a pan-epigenome, ii) evaluate its relationship with the pangenome structure and a possible effect on gene transfer (i.e. HGT) and multireplicon structure, and iii) inspect the occurrence of an uneven distribution of methylated motifs with respect to coding sequences, thus indicating the possibility to relate epigenomic differences to gene regulation. Results obtained showed the presence of a wide pan-epigenome, including both 4mC and 6mA palindromic and nonpalindromic motifs. These motifs have differential occurrence in *S. meliloti* replicons and between coding and noncoding sequences and partially mirroring gene exchange rates. Taken together, our results suggest that the epigenome may impact *S. meliloti* microevolution and adaptation.

Results

S. meliloti has a large pan-epigenome

SMRT sequencing of 21 *S. meliloti* strains yielded genomes with an average coverage of 155X, average genome size of 7.1 Mbp and from 3 to 9 contigs. Contig mapping to *S. meliloti* 1021 genome allowed to identify, for 16 strains, a certain number of contigs not mapping to the sequences of *S. meliloti* 1021 chromosome, pSymB chromid and pSymA megaplasmid. These contigs were labelled as “other” and range from few Kbp (e.g., 3’455 bp in AK75) to hundreds of Kbp (e.g., 328’353 bp for CE480L). We may hypothesize that some of these contigs could correspond to additional plasmids, as already shown for AK83 and SM11 strains (Schneiker-Bekel et al. 2011; Galardini et al. 2011) and as suggested by the number of genes coding for the replication initiation protein RepC identified in the genome sequences (Supplementary Information file, Table S1). Considering the entire panel of strains, a total of 27 methylated raw motifs were identified, further collapsed into 16 different motifs (Supplementary Information file, Table S2), since they shared a common core sequence. Seven motifs contained 4mC, while nine contained 6mA. Ten motifs were palindromic or nearly palindromic. The proportion of motifs methylated ranged from an average of 0.02 to nearly fully methylated (0.99) for GANTC (Supplementary dataset S3). The nearly full methylation for GANTC (0.98-0.99) is in line with the expectation from cells collected in the late exponential growth phase, as previously observed (diCenzo et al. 2022).

Among the identified methylated motifs, two groups can be defined: a “core” set of motifs (shared by all strains analyzed) and a “dispensable” one (occurring in some strains only). The core epigenome was composed of 9 motifs, which are methylated in all strains, while the other 7 motifs formed the dispensable epigenome, being methylated from a minimum of 2 to a maximum of 19 strains. No strain-exclusive motif was found (Figure 1, Supplementary dataset S3). The core GANTC motif is that recognized by the highly conserved cell cycle regulated CcrM methyltransferase and it was expected to be part of the core epigenome (being *ccrM* part of the core genome), as previously reported (diCenzo et al. 2022). We here hypothesize that the other core motifs are due to a common set of DNA-MTases shared between the strains, while the dispensable epigenome suggests the presence of DNA MTases encoded by genes belonging to the accessory genomes. Indeed, genome annotation identified from 4 to 9 DNA MTases in the genomes (Supplementary information file, Table S3, and Supplementary dataset S2), which agrees with the variability in methylated

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

DNA motifs found among strains. When inspecting if strains with differences in the methylation profiles were showing differences in the transfer of genetic material too (possibly due to methylation profiles linked to R-M systems), we found significant differences (up to 2 orders of magnitude) in pairwise gene transfer efficiency in five randomly selected strains (AO641M, BL225C, CE480L, CL374FF and CO438LL) (Supplementary information file, Figure S1 and Tables S4-S6). This result may suggest that differential presence of methylated profiles could have some impact on gene transfer among strains of *S. meliloti*.

Finally, the phylogenetic relationships among these strains have a relatively limited overlap with similarities in their epigenomes. Among strains whose shared epigenome pattern mirrors their phylogenetic relationships we found strains AO641M, AL703GG and CL347FF. These are evolutionarily strongly related and share a high density of methylation on RCCAGCC (Figure 1). In most of the other cases (e.g., SM11), nor core or dispensable genome phylogenies clearly reflect epigenome profiles, overall indicating that both the presence and the frequency of methylation in *S. meliloti* only partially reflects genomic relatedness.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

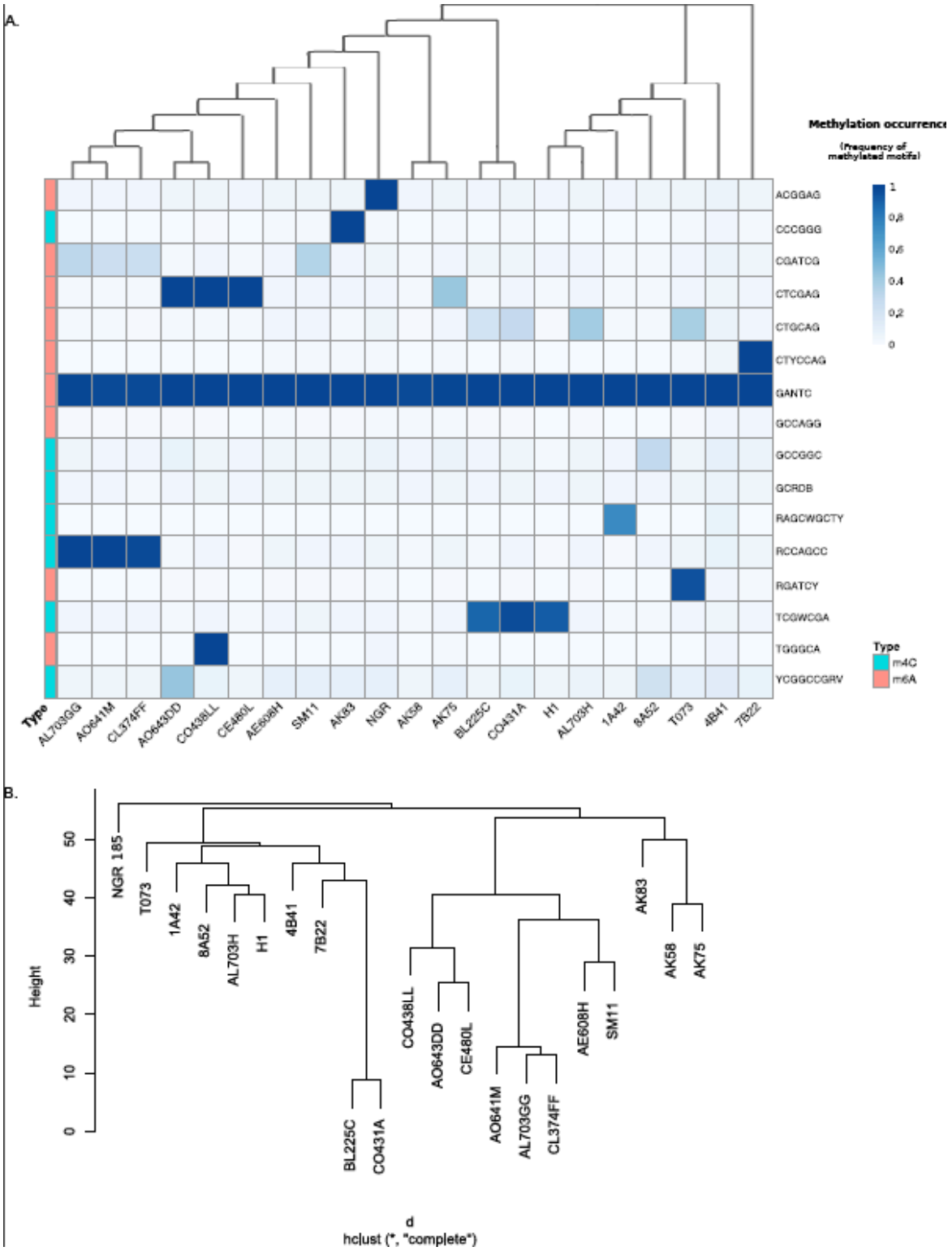


Figure 1. *S. meliloti* pan-epigenome partially reflects strains genomic relatedness. A) Matrix of occurrence of methylated profiles and their level of methylation. Clustering of strains is based on unrooted core gene phylogeny from a concatemer of 1’484’250 nucleotide residues. Clustering of motifs is based on Bray-Curtis distance. B) Clustering of the dispensable genome based on 11’178 genes.

Intergenic and upstream sequences display low methylation values

To check whether the detected methylation profiles could be involved in gene expression control, we inspected the occurrence of methylated motifs in coding, non-coding, intergenic and upstream to genes

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

regions (CDS, nCDS, tIG and US sequences, as defined in Figure S2) in each of the 21 genomes analyzed (Supplementary Dataset S3). In general, tIG sequences showed a statistically significant lower frequency of methylated motifs with respect to CDS, nCDS and US for most of the motifs ($p<0.05$ one-way ANOVA, Dunn's post hoc test, Supplementary Dataset S3). These results indicate that methylation frequency is not evenly distributed across functional genomic features (i.e., CDS, tIG, US). Methylation of specific DNA sequence motifs is known to be related to DNA transcriptional regulation (Sanchez-Romero and Casadesus 2020). For instance, in *Caulobacter crescentus* the binding properties and activation of gene expression by the transcription factor GcrA is dependent on the methylation of the coding DNA strains (Mohapatra et al. 2020). The lower level of methylation in tIG may suggest possible role in transcriptional regulation of some of the found motifs. The biplot of a Principal Component analysis of the methylation frequencies across motifs (Figure 2b) reinforced this hypothesis, clearly showing a genomic feature-dependent clustering and a strong separation of regions likely containing putative regulatory elements (tIG and US) from coding sequences (CDS and nCDS) of the variance vectors. To evaluate the strain-by-strain variation across functional features, the coefficient of variation (CV) in frequency of methylation was computed (Figure 2c). The smallest values of CV, corresponding to the lowest inter-strain variability were found for the highly abundant methylated motifs GANTC and GCRDB, and for GCCAGG (ranging between 10^{-5} and 10^{-3}). The other motifs exhibited a CV of methylated fractions 100 times higher than that of these highly methylated motifs (ranging around 10^{-1}), which may further suggest an involvement of methylation of such motifs in methylation-mediated variability of phenotype/gene expression among strains. However, additional investigation of methylation pattern under different growth conditions coupled with transcriptomic analyses are needed to sustain the hypothesis of an epigenomic cause of phenotypic plasticity and transcriptional variation in *S. meliloti* strains.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

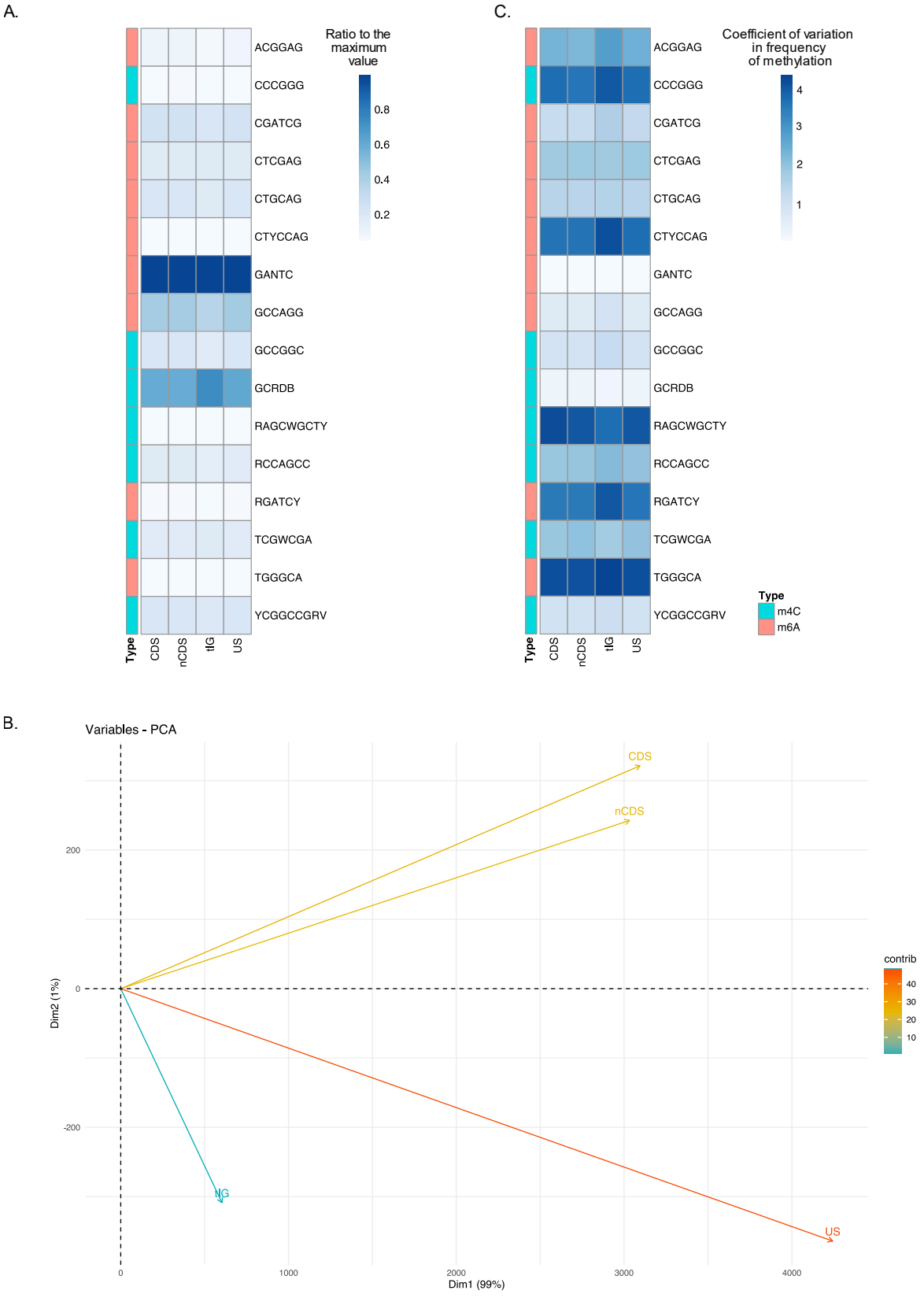
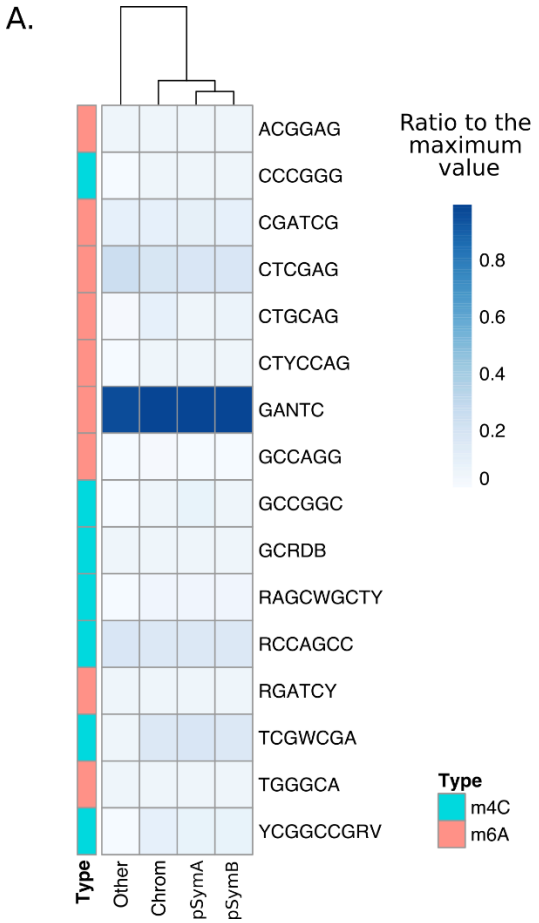


Figure 2. Functional genomic features show different methylation status. A) Heatmap of the proportion of ratio of methylation (ratio to the maximal value) functional features; B) Principal Component Analysis of values of frequency of methylation for tIG, US, CDS and nCDS. A biplot reporting the motifs mostly contributing to differences among features is reported; C) Heatmap of the coefficient of variation in frequency of methylation.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Methylation profiles are mainly similar across replicons

Since *S. meliloti* harbors a multipartite genome structure, which is connected to a functional specialization of the different replicons (chromosome, a chromid and a megaplasmid) and to different evolutionary dynamics (Galardini et al. 2013; diCenzo et al. 2014; diCenzo et al. 2016), we tested the following hypothesis: can DNA methylation profiles reflect such intragenomic differences among replicons? The differential density of methylated motifs along the multipartite *S. meliloti* genome was then analyzed, to check for the presence of a possible uneven occurrence of methylated motifs across the *S. meliloti* replicons (chromosome, pSymB chromid, pSymA megaplasmid and "other") (Figure 3; Supplementary Dataset 3). Out of the 16 motifs identified, 10 showed differences in the frequency of methylation between "other" and the rest of the replicons, 4 showed differences between the chromosome and pSymA, 4 between pSymA and pSymB, and 1 only between the chromosome and pSymB. Interestingly, the methylation frequencies of "other" replicons displayed differences even for GANTC motifs (and particularly low in AK83, 0.81 vs. 0.99 in the rest of the genome), suggesting an uncoupling of the control of DNA duplication of such elements (likely corresponding to plasmids, as indicated by the number of *repC* genes found, Supplementary information file). This possible uncoupling of DNA replication from cell cycle control in the divided *S. meliloti* genome was previously suggested for the megaplasmid pSymA in the terminally differentiated cells (bacteroids) inside root nodules (diCenzo et al. 2022). Moreover, since the "other" replicons are part of the dispensable genome fraction, we cannot exclude that the observed differential frequency of some methylation motifs may reflect their recent HGT. Indeed, it has been shown in rhizobia that genes deriving from recent horizontal gene transfer (HGT) events can be subjected to transcriptional silencing (Shi et al. 2022).



B.

	Chromosome	pSymB	pSymA	other
Chromosome				
pSymB	1			
pSymA	4	4		
other	5	2	3	

Figure 3. Relative frequency of methylated motifs across the multipartite replicons of *S. meliloti*. A) Heatmap reporting the ratios to maximum values for each motif. B) Number of methylated motifs with statistically significant differences among replicons.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Discussion

Genome-wide DNA methylation patterns are key elements in the genotype-to-phenotypes translations, having profound consequences on adaptation and phenotypic plasticity (Herman and Sultan 2016). In prokaryotes, DNA methylation can have several roles, from DNA repair mechanisms, cell cycle control, self-DNA recognition and transcriptional regulation (Oliveira and Fang 2021). Mirroring these multiple roles of DNA methylation, the phylogeny of DNA methyltransferases is also complex, including both vertical and horizontal transmission (Harris and Goldman 2020). This, in turn, may give rise to both DNA methylation patterns shared among strains and to strain-specific DNA methylation profiles, resulting from genes harbored in the core and in the dispensable genome fraction, respectively. In agreement with such hypotheses, evidence from our data on *S. meliloti* indicates that a shared by all (core) and an unshared (dispensable) set of DNA methylation profiles exist. Such epigenomic landscape, with a core and dispensable epigenome, is in line with the pangenomic structure of this species (Galardini et al. 2013). However, the phylogenetic trees on core and dispensable genomes of the *S. meliloti* strains only partially mirror epigenome similarities, suggesting that epigenome profiling (presence and frequency of methylated DNA motifs) cannot be simply explained by evolutionary relatedness. Indeed, considering the possibility that a substantial fraction of the methylated motifs may derive from HGT events (being the motifs occurring in a part only of the strains) we should consider the differential influence acquired DNA methylation profiles may have on the following acquisitions of DNA. When DNA methylation profiles are linked to R-M systems, their acquisition strongly reduces the acquisition of novel DNA (and of potentially novel DNA methyltransferase and consequently new DNA methylation profiles) (Jeltsch 2003). This may bring to a reduction of gene flow among strains with different methylation profiles, favoring the maintenance of cohesive population structures, which in turn can give rise to a new lineage in the population (Oliveira et al. 2016). On the contrary, acquired DNA methylation profiles unlinked to the restriction systems have potentially no effect over novel acquisitions. Consequently, we can figure out that the pattern of the dispensable (viz. unshared) epigenome is the result of the timing of acquisition of “neutral” and “negative” (to the HGT) profiles, corresponding to orphan DNA MTases and R-M systems, respectively. This model will lead to a mirroring of the dispensable genome phylogeny only for those branches where mostly “neutral” profiles have been acquired. The experimental evidence of differences among strains for gene transfer frequencies in our study may suggest the presence of some “negative”-to-HGT methylation profiles. However, from our data we cannot rule out which profile can be more related to differences in gene transfer among the selected strains.

Still concerning the relationships between HGT and DNA methylation, we inspected the multi-replicon structure of *S. meliloti* genome with respect to the pattern of methylation of the chromosome, the chromid pSymB and the symbiotic megaplasmid pSymA, the three resident replicons of *S. meliloti* (Galibert et al. 2001). Evolutionary analyses on *S. meliloti* genomes clearly indicated that pSymA megaplasmid is an alien element which was recently acquired by the resident genome, while pSymB is a chromid, which stabilizes in the ancestral *S. meliloti* genome (diCenzo et al. 2014; Galardini et al. 2013, Galibeert et al. 2001). In line with this timing of replicon acquisition, statistically significant differences between the chromosome and the other replicons were limited for pSymB (1 motif only), but higher for pSymA (4 motifs) and “other” (7 motifs). As expected, differences between pSymA and pSymB were relevant and in line with those between pSymA and the chromosome (4 motifs). This evidence possibly indicates that the DNA methylation status is another parameter which can allow to define genomic integration in multipartite genomes, as differences between chromosome, chromid and plasmids. In line with such results, there are recent indication that HGT elements (i.e., plasmids) show avoidance of target of R-M systems, which may be reflected in different methylation abundance for specific target motifs (Shi et al. 2022)). Moreover, DNA regions originating from HGT, in general tend to have low expression levels and slowly integrate into the transcriptional network (Lercher and Pál 2008). This low level of expression can be related to the presence of xenogeneic silencers, as MucR in *Sinorhizobium fredii* (Jiao et al. 2022), which preferentially targets regions with high AT%. We may speculate that the different methylation frequency for some motifs of “other” replicons may partially reflect a stronger control of gene expression operated during the growth conditions applied.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A hypothesis on the role of the DNA methylation on gene expression regulation was drawn from the results obtained when measuring the frequency of methylation in upstream, intergenic and coding sequences where intergenic regions showed for some motifs lower values of methylation frequencies compared to coding sequences, as well as a high strain-by-strain variability. Such results in fact can be interpreted in terms of involvement of DNA methylation in gene expression control operated at the promoter/intergenic level by differential binding/accessibility of transcription factors. Large strain-by-strain variation in transcriptomes of *S. meliloti* strains has been previously shown and related to genotypic interaction with host plant (Fagorzi et al. 2021). We cannot a priori exclude that strain-by-strain epigenomic differences can have an influence over the observed transcriptomic variations.

In conclusion, here we showed for the first time that, in bacteria, a large pan-epigenome can exist. Moreover, using *S. meliloti* (a nitrogen fixing, multipartite genome model organism), while highlighting that replicons share a relatively similar methylation profiles, we recognized that upstream and intergenic regions can be less methylated than coding sequences and that this pattern of differential methylation is highly variable among strains. These data allow to sustain the hypothesis that DNA methylation may play a role in gene regulation and in phenotypic variation at the strain level. Investigation on DNA MTases mutants and additional analyses of genome-wide methylation in different culture conditions are needed for thoroughly testing this hypothesis.

Materials and Methods

S. meliloti strains, growth conditions and DNA extraction

S. meliloti strains used in this work are listed in Table S7 (Supplementary Information File). All strains were routinely grown on TY medium. Overnight cultures of all strains grown in 10 ml TY medium overnight at 30˚C with shaking (130 rpm). To obtain late exponential/stationary phase samples, cultures were harvested after 24 hours of growth at OD_{600nm} values of ~ 1.4. In all cases, cultures were streaked on TY plates to check for contamination and then centrifuged (8,200 g, 10 minutes, 4°C); the full 10 mL was centrifuged and transferred to a 2 mL tube. DNA was isolated using Qiagen PowerLyser PowerSoil Pro KIT according to manufacturer instructions.

SMRT genome sequencing of S. meliloti strains

Single-Molecule Real Time (SMRT) sequencing was performed in house at the University of Florence using Pacific Biosciences Sequel instrument (Eid et al. 2009). Genomic DNA was sheared to 20 Kbp using g-TUBEs (Covaris Inc., Woburn, MA, USA). Sheared DNA was treated with exonuclease to remove single stranded ends and DNA damage repair mix followed by end repair and ligation of barcoded blunt adapters using SMRTbell Template Prep Kit 2.0 (PacBio, Menlo Park, CA, USA). Libraries were purified with AMPure PB beads (Beckman Coulter, Brea, CA, USA) and eight libraries with different barcodes were pooled at equimolar ratios and purified with AMPure PB beads. SMRTbell template libraries were prepared using a Sequel Binding Kit 3.0 (PacBio, Menlo Park, CA, USA), and sequenced on a Sequel instrument using a v3 or v4 sequencing primer, 1M v3 SMRT cells, and Version 3.0 sequencing chemistry.

Genome assembly was performed using the PacBio SMRT Link software (PacBio, Menlo Park, CA, USA). Briefly, raw reads were filtered using SFilter to remove short reads and reads derived from sequencing adapters. Filtered reads were assembled into contig though the Microbial Assembly SMRT Link tool. Each contig has been aligned with the genome of *S. meliloti* 1021 (GenBank assembly accession GCA_000006965.1) using Mauve software (Darling et al. 2004) and categorized into Chromosome, pSymA-like replicon, pSymB-like replicon and “other” (when no mapping to *S. meliloti* 1021 was found) (Supplementary dataset S1). Genomes were reannotated with Prokka (V 1.12-beta). Genome sequences are deposited under NCBI BioProject Accession PRJNA681719.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Identification of methylated bases and mapping to genomic features

The assembled genomes were used as reference to map modified sites identified through kinetic analysis of the aligned DNA sequence data with the SMRT Link software ver.8.0.0.80529 (Pacific Biosciences, Menlo Park, CA, USA) (Flusberg et al. 2010), using default options; the number of mapped bases per sample is provided in Supplementary dataset S1. Modified sites were then grouped into motifs using MotifFinder. Downstream analyses were performed using MeStudio (Riccardi et al. 2022). MeStudio input files are the genome sequences in FASTA format, the genome annotation in GFF3 format and the PacBio-derived output file with methylation annotations in GFF3 format. MeStudio allows to determine the motifs distribution along: i) protein-coding gene with accordant (sense) strand (CDS), ii) regions that fall between annotated genes (true intergenic, tIG), iv) regions upstream to the reading frame of a gene, with accordant strand (US) (Supplementary information file, Figure S2). The differences in the frequencies of methylated motifs among categories were evaluated by post hoc Dunn's test and Pearson correlation.

Phylogenetic reconstruction

To construct an unrooted core gene phylogeny, the pangenome of the 21 *S. meliloti* strains was calculated using Roary (version 3.11.2) (Page et al. 2015) with a percent identity threshold of 95% (Supplementary dataset S2). The nucleotide sequences of the 4,633 core genes (identified as those found in at least 99% of the genomes; Supplementary dataset S2) were aligned with PRANK (Löytynoja 2014) and the alignments were trimmed and concatenated. The concatenated alignment was used to construct a maximum likelihood phylogeny (the bootstrap best tree following 100 bootstrap replicates) using RAxML 8.4.2 (Stamatakis 2006) with the GTRCAT model. Phylogeny was visualized with the online iTOL webserver (Letunic and Bork 2021).The dispensable gene set of the 21 strains was used to create a presence/absence matrix. Hierarchical clustering was performed in R (R Core Team 2014) with *hclust* function based on Euclidean distance, with "complete" agglomeration method.

Acknowledgement

This work was partially supported by the grant MICRO4Legumes, D.M.n.89267 (Italian Ministry of Agriculture) to A.M. I.P. is supported by a PhD fellowship from PNRR, D.M. n. 351/2022 (Italian Ministry of University and Research). L.C. is supported by a PhD fellowship from MICRO4Legumes. C.F. is supported by a post-doctoral fellowship from the LEGU-MED project (PRIMA Foundation) Call 2019. We are grateful to George C. diCenzo for useful insights and suggestions during data analysis and manuscript writing.

References

Atack JM, Yang Y, Seib KL, Zhou Y, Jennings MP. 2018. A survey of Type III restriction-modification systems reveals numerous, novel epigenetic regulators controlling phase-variable regulons; phasevarions. *Nucleic Acids Res*.

Batstone RT, Burghardt LT, Heath KD. 2022. Phenotypic and genomic signatures of interspecies cooperation and conflict in naturally occurring isolates of a model plant symbiont. *Proceedings of the Royal Society B: Biological Sciences* 289.

Batstone RT, O'Brien AM, Harrison TL, Frederickson ME. 2020. Experimental evolution makes microbes more cooperative with their local host genotype. *Science (1979)*.

Beaulaurier J, Schadt EE, Fang G. 2019. Deciphering bacterial epigenomes using modern sequencing technologies. *Nat Rev Genet*.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Bellabarba A, Bacci G, Decorosi F, Aun E, Azzarello E, Remm M, Giovannetti L, Viti C, Mengoni A, Pini F. 2021. Competitiveness for Nodule Colonization in Sinorhizobium meliloti: Combined In Vitro -Tagged Strain Competition and Genome-Wide Association Analysis . *mSystems*.

Bertalan M, Albano R, de Pádua V, Rouws L, Rojas C, Hemerly A, Teixeira K, Schwab S, Araujo J, Oliveira A, et al. 2009. Complete genome sequence of the sugarcane nitrogen-fixing endophyte Gluconacetobacter diazotrophicus Pal5. *BMC Genomics*.

Biondi EG, Tatti E, Comparini D, Giuntini E, Mocali S, Giovannetti L, Bazzicalupo M, Mengoni A, Viti C. 2009. Metabolic capacity of Sinorhizobium (Ensifer) meliloti strains as determined by phenotype microarray analysis. *Appl Environ Microbiol*.

Blow MJ, Clark TA, Daum CG, Deutschbauer AM, Fomenkov A, Fries R, Froula J, Kang DD, Malmstrom RR, Morgan RD, et al. 2016. The Epigenomic Landscape of Prokaryotes. *PLoS Genet*.

Cangioli L, Checcucci A, Mengoni A, Fagorzi C. 2021. Legume tasters: Symbiotic rhizobia host preference and smart inoculant formulations. *Biological Communications*.

Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res*.

diCenzo GC, Cangioli L, Nicoud Q, Cheng JHT, Blow MJ, Shapiro N, Woyke T, Biondi EG, Alunni B, Mengoni A, et al. 2022. DNA Methylation in Ensifer Species during Free-Living Growth and during Nitrogen-Fixing Symbiosis with Medicago spp. . *mSystems*.

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

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

diCenzo GC, Finan TM. 2017. The Divided Bacterial Genome: Structure, Function, and Evolution. *Microbiology and Molecular Biology Reviews*.

diCenzo GC, MacLean AM, Milunovic B, Golding GB, Finan TM. 2014. Examination of Prokaryotic Multipartite Genome Evolution through Experimental Genome Reduction. *PLoS Genet*.

Dicenzo GC, Mengoni A, Perrin E. 2019. Chromids Aid Genome Expansion and Functional Diversification in the Family Burkholderiaceae. *Mol Biol Evol*.

Dicenzo GC, Zamani M, Checcucci A, Fondi M, Griffitts JS, Finan TM, Mengoni A. 2019. Multidisciplinary approaches for studying rhizobium–legume symbioses. *Can J Microbiol*.

Dordet-Frisoni E, Vandecasteele C, Contarin R, Sagné E, Baranowski E, Klopp C, Nouvel LX, Citti C. 2022. Impacts of Mycoplasma agalactiae restriction-modification systems on pan-epigenome dynamics and genome plasticity. *Microb Genom* 8:1–23.

Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, et al. 2009. Real-time DNA sequencing from single polymerase molecules. *Science (1979)*.

Epstein B, Burghardt LT, Heath KD, Grillo MA, Kostanecki A, Hämälä T, Young ND, Tiffin P. 2022. Combining GWAS and population genomic analyses to characterize coevolution in a legume-rhizobia symbiosis. *Mol Ecol*:1–14.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Estibariz I, Ailloud F, Woltemate S, Bunk B, Spröer C, Overmann J, Aebischer T, Meyer TF, Josenhans C, Suerbaum S. 2020. In vivo genome and methylome adaptation of CAG-negative helicobacter pylori during experimental human infection. *mBio*.

Fagorzi C, Bacci G, Huang R, Cangioli L, Checcucci A, Fini M, Perrin E, Natali C, diCenzo GC, Mengoni A. 2021. Nonadditive transcriptomic signatures of genotype-by-genotype interactions during the initiation of plant-rhizobium symbiosis. *mSystems* 6.

Ferri L, Gori A, Biondi EG, Mengoni A, Bazzicalupo M. 2010. Plasmid electroporation of Sinorhizobium strains: The role of the restriction gene hsdR in type strain Rm1021. *Plasmid*.

Fioravanti A, Fumeaux C, Mohapatra SS, Bompard C, Brilli M, Frandi A, Castric V, Villeret V, Viollier PH, Biondi EG. 2013. DNA Binding of the Cell Cycle Transcriptional Regulator GcrA Depends on N6-Adenosine Methylation in Caulobacter crescentus and Other Alphaproteobacteria. *PLoS Genet*.

Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, Korlach J, Turner SW. 2010. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods*.

Galardini M, Mengoni A, Brilli M, Pini F, Fioravanti A, Lucas S, Lapidus A, Cheng JF, Goodwin L, Pitluck S, et al. 2011. Exploring the symbiotic pangenome of the nitrogen-fixing bacterium Sinorhizobium meliloti. *BMC Genomics*.

Galardini M, Pini F, Bazzicalupo M, Biondi Emanuele G., Mengoni A. 2013. Replicon-dependent bacterial genome evolution: The case of Sinorhizobium meliloti. *Genome Biol Evol*.

Galardini M, Pini F, Bazzicalupo M, Biondi Emanuele G, Mengoni A. 2013. Replicon-Dependent Bacterial Genome Evolution : The Case of. *Mol Biol* 5:542–558.

Galibert F, Finan TM, Long SR, Pühler A, Abola P, Ampe F, Barloy-Hubler F, Barnet MJ, Becker A, Boistard P, et al. 2001. The composite genome of the legume symbiont Sinorhizobium meliloti. *Science (1979)*.

Harris AJ, Goldman AD. 2020. The complex phylogenetic relationships of a 4mC/6mA DNA methyltransferase in prokaryotes. *Mol Phylogenet Evol*.

Herman JJ, Sultan SE. 2016. DNA methylation mediates genetic variation for adaptive transgenerational plasticity. *Proceedings of the Royal Society B: Biological Sciences* 283.

Jeltsch A. 2003. Maintenance of species identity and controlling speciation of bacteria: A new function for restriction/modification systems? In: Gene.

Jiao J, Zhang B, Li ML, Zhang Z, Tian CF. 2022. The zinc-finger bearing xenogeneic silencer MucR in α-proteobacteria balances adaptation and regulatory integrity. *ISME Journal*.

Jones PA. 2012. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat Rev Genet*.

Kahramanoglou C, Prieto AI, Khedkar S, Haase B, Gupta A, Benes V, Fraser GM, Luscombe NM, Seshasayee ASN. 2012. Genomics of DNA cytosine methylation in Escherichia coli reveals its role in stationary phase transcription. *Nat Commun*.

Kumar S, Karmakar BC, Nagarajan D, Mukhopadhyay AK, Morgan RD, Rao DN. 2018. N4-cytosine DNA methylation regulates transcription and pathogenesis in Helicobacter pylori. *Nucleic Acids Res*.

Ledermann R, Schulte CCM, Poole PS. 2021. How rhizobia adapt to the nodule environment. *J Bacteriol*.

Lercher MJ, Pál C. 2008. Integration of horizontally transferred genes into regulatory interaction networks takes many million years. *Mol Biol Evol*.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Letunic I, Bork P. 2021. Interactive tree of life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res*.

Löytynoja A. 2014. Phylogeny-aware alignment with PRANK. *Methods in Molecular Biology*.

Militello KT, Mandarano AH, Varechtchouk O, Simon RD. 2014. Cytosine DNA methylation influences drug resistance in Escherichia coli through increased sugE expression. *FEMS Microbiol Lett*.

Mohapatra SS, Fioravanti A, Vandame P, Spriet C, Pini F, Bompard C, Blossey R, Valette O, Biondi EG. 2020. Methylation-dependent transcriptional regulation of crescentin gene (creS) by GcrA in Caulobacter crescentus. *Mol Microbiol* 114:127–139.

Oliveira PH. 2021. Bacterial Epigenomics: Coming of Age. *mSystems*.

Oliveira PH, Fang G. 2021. Conserved DNA Methyltransferases: A Window into Fundamental Mechanisms of Epigenetic Regulation in Bacteria. *Trends Microbiol*.

Oliveira PH, Touchon M, Rocha EPC. 2016. Regulation of genetic flux between bacteria by restriction-modification systems. *Proc Natl Acad Sci U S A*.

Page AJ, Cummins CA, Hunt M, et al. 2015. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31.22:3691–3693.

Pini F, De Nisco NJ, Ferri L, Penterman J, Fioravanti A, Brilli M, Mengoni A, Bazzicalupo M, Viollier PH, Walker GC, et al. 2015. Cell Cycle Control by the Master Regulator CtrA in Sinorhizobium meliloti. *PLoS Genet*.

Poole P, Ramachandran V, Terpolilli J. 2018. Rhizobia: From saprophytes to endosymbionts. *Nat Rev Microbiol*.

R Core Team. 2014. R Core Team (2014). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.*

Reva ON, Swanevelder DZH, Mwita LA, Mwakilili AD, Muzondiwa D, Joubert M, Chan WY, Lutz S, Ahrens CH, Avdeeva L V., et al. 2019. Genetic, Epigenetic and Phenotypic Diversity of Four Bacillus velezensis Strains Used for Plant Protection or as Probiotics. *Front Microbiol*.

Riccardi C, Passeri I, Cangioli L, Fagorzi C, Mengoni A, Fondi M. 2022. MeStudio : crossing methylation and genomic features for comparative epigenomic analyses. :1–10.

Sánchez-Romero MA, Casadesús J. 2020. The bacterial epigenome. *Nat Rev Microbiol*.

Shaw LP, Rocha EPC, MacLean RC. 2022. Restriction-modification systems have shaped the evolution and distribution of plasmids across bacteria. *bioRxiv* [Internet]:2022.12.15.520556. Available from: https://www.biorxiv.org/content/10.1101/2022.12.15.520556v1

Shi WT, Zhang B, Li ML, Liu KH, Jiao J, Tian CF. 2022. The convergent xenogeneic silencer MucR predisposes α-proteobacteria to integrate AT-rich symbiosis genes. *Nucleic Acids Res* 50:8580–8598.

Stamatakis A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*.

Touchon M, Hoede C, Tenaillon O, Barbe V, Baeriswyl S, Bidet P, Bingen E, Bonacorsi S, Bouchier C, Bouvet O, et al. 2009. Organised genome dynamics in the Escherichia coli species results in highly diverse adaptive paths. *PLoS Genet*.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.31.542881; this version posted June 1, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Vasu K, Nagaraja V. 2013. Diverse Functions of Restriction-Modification Systems in Addition to Cellular Defense. *Microbiology and Molecular Biology Reviews*.

Yuan W, Zhang Y, Riaz L, Yang Q, Du B, Wang R. 2021. Multiple antibiotic resistance and DNA methylation in Enterobacteriaceae isolates from different environments. *J Hazard Mater*.

Zweiger G, Marczynski G, Shapiro L. 1994. A Caulobacter DNA methyltransferase that functions only in the predivisional cell. *J Mol Biol*.

plants

MDPI

Article

Effect of Wastewater on the Composition of Bacterial Microbiota of *Phragmites australis* Used in Constructed Wetlands for Phytodepuration

Lisa Cangioli [1,†], Maria Salobehaj [2,†], Sara Del Duca [1], Camilla Fagorzi [1], Chiara Berardi [3], Ester Coppini [3], Donatella Fibbi [3], Renato Fani [1,*] and Alberto Vassallo [4,*]

[1] Department of Biology, University of Florence, 50019 Sesto Fiorentino, Italy
[2] Center for Magnetic Resonance (CERM), 50019 Sesto Fiorentino, Italy
[3] Gestione Impianti di Depurazione Acque (G.I.D.A.) SpA, 59100 Prato, Italy
[4] School of Biosciences and Veterinary Medicine, University of Camerino, 62032 Camerino, Italy
* Correspondence: renato.fani@unifi.it (R.F.); alberto.vassallo@unicam.it (A.V.)
† These authors contributed equally to this work.

Abstract: Phytodepuration occurs in the plant-mediated remediation processes exploited to remove pollutants from wastewater, and *Phragmites australis* is one of the most used plants. This goal is achieved using constructed wetlands (CW), which are engineered systems designed to mimic the natural processes of pollutants removal. The aim of this work was to characterize the bacterial communities associated to *P. australis*, soils, and permeates of the CW of Calice (Prato, Italy), to evaluate the possible effect of wastewaters on the CW bacterial communities, through a next-generation sequencing-based approach. A total of 122 samples were collected from different tissues of *P. australis* (i.e., roots, aerial parts, and stem), soil (i.e., rhizospheric and bulk soil), and permeates, and analyzed. All samples were collected during five sampling campaigns, with the first one performed before the activation of the plant. Obtained results highlighted a specific microbiota of *P. australis*, conserved among the different plant tissues and during time, showing a lower alpha diversity than the other samples and not influenced by the more complex and variable environmental (soils and permeates) bacterial communities. These data suggest that *P. australis* is able to select and maintain a defined microbiota, a capacity that could allow the plant to survive in hostile environments, such as that of CW.

Keywords: stable plant microbiota; landfill leachate; soil microbiota; wastewater treatment plant; next-generation sequencing; *Phragmites australis*; phytoremediation; constructed wetlands

check for updates

Citation: Cangioli, L.; Salobehaj, M.; Del Duca, S.; Fagorzi, C.; Berardi, C.; Coppini, E.; Fibbi, D.; Fani, R.; Vassallo, A. Effect of Wastewater on the Composition of Bacterial Microbiota of *Phragmites australis* Used in Constructed Wetlands for Phytodepuration. *Plants* **2022**, *11*, 3210. https://doi.org/10.3390/plants11233210

Academic Editor: Robert Czajkowski

Received: 30 September 2022
Accepted: 21 November 2022
Published: 23 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The term phytodepuration indicates the plant-mediated remediation processes exploited to remove pollutants from water and wastewater [1]. This goal is achieved using constructed wetlands (CW), which are engineered systems designed to mimic the natural processes occurring in natural wetlands, although in a more controlled environment, to remove organic compounds, suspended solids, microbes, and heavy metals [2]. Moreover, as recently reviewed, CW can also limit the diffusion of antibiotics, antimicrobial resistance genes (ARGs), and antibiotic resistant bacteria (ARB) [1]; these are listed among the so-called "contaminants of emerging concern" (CEC), and their presence in wastewater may represent a serious danger as they can escape typical controls, spread in the environment, and even evolve [3]. Indeed, the new term "evolving CEC" (e-CEC) has been proposed for ARGs and ARB [3]. CW are recognized as a reliable technology for wastewater treatment, as well as being sustainable and cost-effective [4]; in fact, their use for the treatment of different types of wastewater has been reported, such as those from municipal, farm, and industrial sources [5]. In addition, after their use in CW, plants can potentially be used for

other purposes within a circular economy framework, as their biomass results enriched of nutrients [6]: for instance, these plants could be used, at least in certain conditions, for the production of animal feedstock, fertilizers, biofuel, biochar, adsorbent, and biogas [6,7].

CW are typically classified basing on their designing parameters, such as hydrology (i.e., open water-surface flow and subsurface flow), kind of macrophyte plant growth (i.e., emergent, submerged, and free-floating), and flow path (i.e., horizontal and vertical) [2].

Plants can remove pollutants in CW through different mechanisms, such as phytoextraction (i.e., root uptake of contaminants with their accumulation in the aerial parts), phytovolatilization (i.e., pollutant uptake followed by translocation and release through the surface of stem and leaves), phytostabilization/phytoimmobilization (i.e., conversion of toxic compounds into non-toxic or less toxic forms having a reduced bioavailability), and phytodegradation (i.e., degradation of organic compounds due to enzymes contained in plant exudates) [8].

Microbial communities represent important components of CW, as they play a crucial role in the removal of pollutants. Indeed, microbes take part in the remediation process directly: for example, bacteria can help the biosorption of heavy metals by increasing their solubility and bioavailability through the reduction of soil pH and the release of chelators, and they can degrade organic compounds [8]. On the other hand, microorganisms, especially rhizospheric ones, also increase plant performances through the improvement of plant growth; for instance, they can stimulate root proliferation, protect against phytopathogens, lower the production of ethylene with their 1-aminocyclopropane-1-carboxylate (ACC) deaminases, fix nitrogen, and/or produce phytohormones such as auxin [9,10].

Phragmites australis (the common reed) is an aquatic emergent plant and it is one of the most used plants in phytodepuration, due to its (i) ability of removing nutrients from different kinds of wastewater, (ii) flood-tolerance, and (iii) resistance against both pollution and (iv) salinity [11,12]. Moreover, it has a wide geographical distribution, and it is particularly present in Europe and Asia [11,12].

P. australis is also used for phytodepuration in the pilot CW at the depuration plant in Calice (Prato, Tuscany, Italy), managed by G.I.D.A. SpA. These CW are designed for the tertiary treatment of landfill leachate (LFL) and are located downstream to a membrane bio-reactor (MBR) pretreating LFLs before their discharge in the main line of a full-scale wastewater treatment plant (WWTP) [13]. Each of the two lines of CW in Calice is a two-stage subsurface flow system (SFS), composed by a vertical system (SFS-v) followed by a horizontal one (SFS-h).

In a previous study, the bacterial culturable community associated to the roots of *P. australis* from the CW of Calice was investigated, covering a time span of 22 months [14]; one of the most interesting aspects of that study was that samples were collected both before and after the activation of the CW, allowing the analysis of the effect of wastewater influx on root-associated bacteria. This investigation highlighted how wastewater exerted a selective pressure leading to a progressive reshaping of bacterial community (e.g., 4 genera present before the activation of CW disappeared and 27 new ones colonized *P. australis* roots upon influx of wastewater). Moreover, time of exposure to wastewater was directly correlated with increasing numbers of bacterial isolates resistant to antibiotics and metal(loid)s.

The present work expands and integrates the previous one, as samples analyzed here were collected at the same timepoints [14]. While in the former work only culturable bacteria were isolated and identified, the aim here is to characterize, applying a culture-independent method, the entire bacterial communities associated to different compartments of *P. australis*, soils of CW, and permeates, through a next-generation sequencing-based investigation. In this regard, despite other similar works already available in literature, this work spans a longer time and takes in consideration not only the roots and rhizosphere of *P. australis*, which are the most studied compartments in CW, but also other anatomical compartments and soil samples collected in all the different parts of CW.

2. Results

2.1. Bacterial Diversity in Plants, Soil and Permeates Microbiome

In this work, a total of 122 samples collected from different plant compartments of *P. australis* (i.e., roots, aerial parts, and stem), soil (i.e., rhizospheric and bulk soil), and permeates were analyzed. All samples were collected during five sampling campaigns from both the vertical and the horizontal systems of CW, except for permeate samples which were not collected for the first sampling, performed before the CW activation.

DNA sequencing of V3-V4 region of 16S rDNA of these samples yielded a total of 17,736,516 reads, and 11,687,785 passed the quality filter (66% of total reads). Reads were assigned to a total of 57,921 Amplicon Sequencing Variants (ASVs) after a clustering step (Dataset S1); ASVs assigned to Archaea (25 ASVs) and ASVs not assigned to any phylum (15,076 ASVs) were removed from the following analysis.

As shown in Figure S1, all rarefaction curves reached a plateau, thus demonstrating a good representation of the microbial communities and a sufficiently deep sampling. Bacterial diversity of all samples was preserved, as shown by the Good's coverage estimator reported in Table S1, and alpha diversity indices were estimated (Table S2).

Concerning the different samples, as expected, above-ground plant compartments (aerial part and stem) showed a reduced alpha diversity in comparison to soil (rhizosphere and bulk soil) and permeates, while root samples showed an intermediate bacterial diversity (Figure 1).

In particular, when compared, only (i) aerial part and stem, and (ii) rhizosphere, bulk soil, and permeates showed a non-significant difference, suggesting a similarity in the bacterial diversity of these samples. In contrast, all the other tested combinations resulted in a significant difference in alpha diversity (p-value < 0.001) (Table S3).

Concerning the different sampling times, (i) aerial part, stem, and root, and (ii) rhizosphere and bulk soil samples were joined in the two groups named (i) plant tissues and (ii) soils, respectively. Plant tissues differed in alpha diversity in relation to the time of sampling; in particular, the fourth sampling was different from the third and the fifth ones (p-value < 0.05), showing a reduction in diversity. The same difference was also observed for permeate samples, which differentiated across times according to Simpson index, indicating that the dominance of bacterial groups changes among samplings (Table S3 and Figure 1).

To explore species diversity among samples, beta diversity was analyzed, performing a non-metric Multidimensional Scaling (NMDS) (Figure 2).

Beta diversity highlighted taxonomic similarities among plant samples (aerial part, stem, root). During this time, the fourth sampling was the most different one, showing a significant p-value (<0.05) when compared with the first, second, and third ones (Table S4). On the contrary, soil and permeate bacterial communities were highly different from each other and from plant tissues samples, and deeply influenced by the time of sampling, as confirmed by the PERMANOVA analysis reported in Table S4. In contrast, no differences were observed considering the two CW systems.

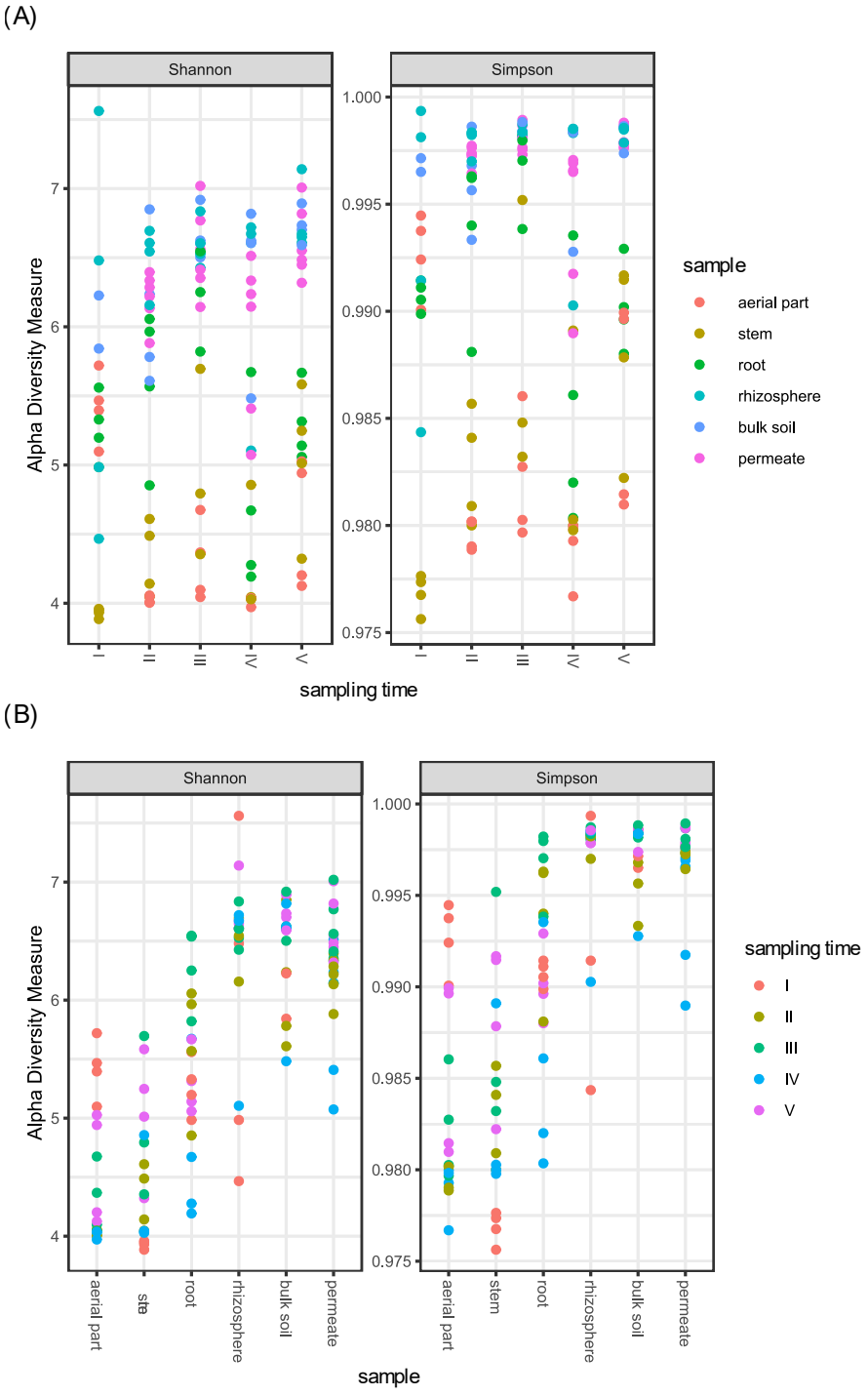


Figure 1. Shannon and Simpson diversity indices of the bacterial communities associated to *P. australis*, soil, and permeate samples collected in the CW during five samplings. (**A**) Samples are divided per sampling time. (**B**) Samples are divided per compartments.

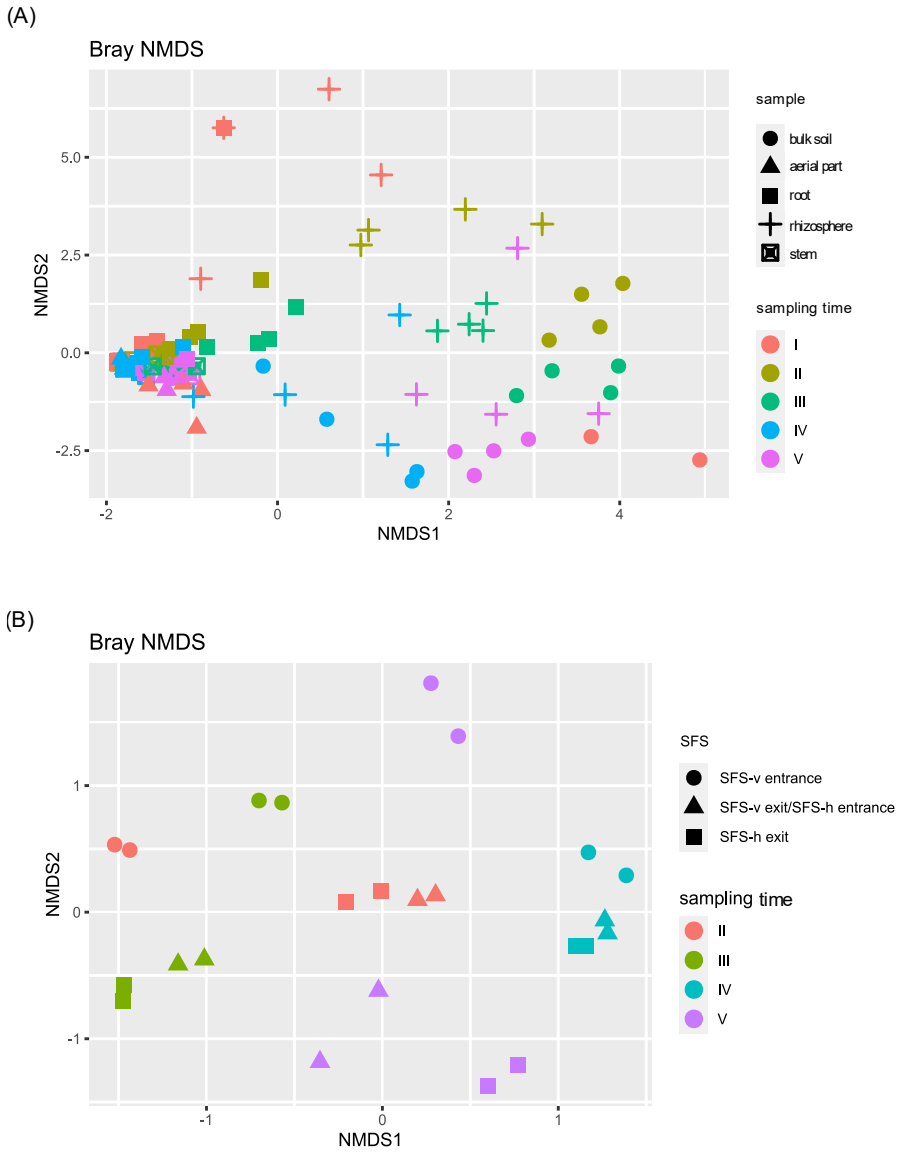


Figure 2. NMDS beta diversity of the bacterial communities of *P. australis*, soil and permeate samples. Samples were divided on the basis of the sampling site of the CW plant: (**A**) those collected directly from the SFS (plant tissues and soils) and (**B**) those collected from the CW irroration system (permeates).

2.2. Taxonomic Composition of Plant, Soil, and Permeate Bacterial Communities

Total ASVs were classified into 1019 genera belonging to 50 phyla, revealing that more than 80% of the ASVs were classified into six phyla, i.e., Proteobacteria (36.78%), Bacteroidota (17.77%), Actinobacteriota (7.49%), Patescibacteria (7.29%), Planctomycetota (5.82%), and Verrucomicrobiota (5.6%). The lowest number of phyla was detected in the aerial part (15 phyla) and stem (20) of the plants, while the highest number was detected in permeate samples, in which 46 different phyla were observed. In all the compartments, the most frequent phyla were Proteobacteria (comprised between 31.6% and 54.2%) and Bacteroidota (between 13.2% and 24.9%). The third most frequent detected phylum was Firmicutes for the aerial part and the stem of *P. australis* (with frequencies of 10% and 7.8%, respectively), Planctomycetota for the roots (6.9%), Patescibacteria for the rhizosphere (8.1%), Actinobacteriota for the bulk soil (12.2%), and Patescibacteria for the permeates (9.8%).

The relative abundances of detected phyla are reported in Figure 3, where they appear to be more similar with one another among the plant tissues (i.e., aerial part, stem, and

root), showing higher relative abundances for Proteobacteria and Cyanobacteria. On the contrary, soil and permeate samples showed higher levels of diversity, in agreement with results regarding alpha diversity reported in Section 2.1.

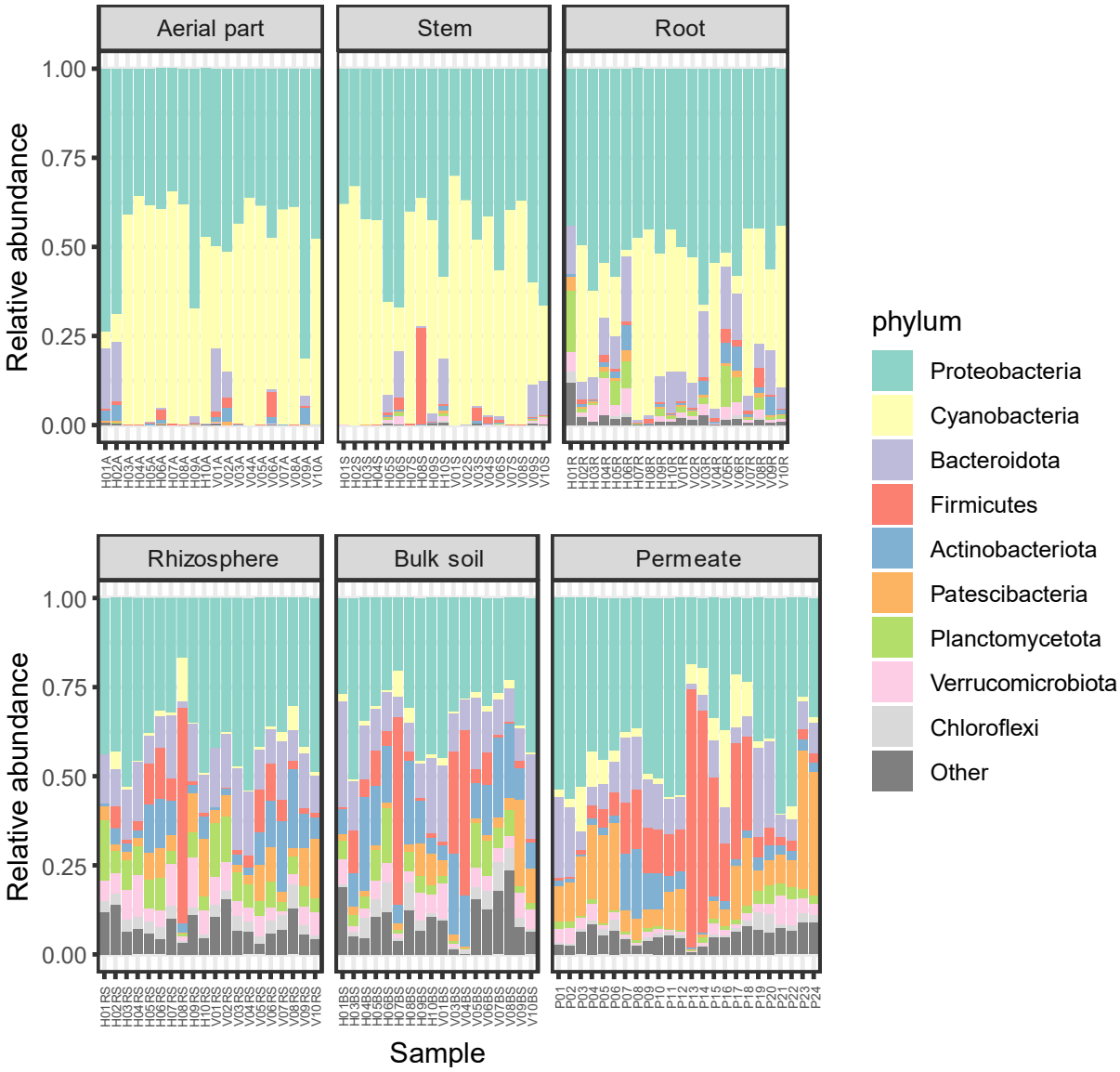


Figure 3. Relative abundances of bacterial phyla in each sample. ASVs representing < 5% of the whole community have been reported as "Other".

2.3. Effect of the Different Sampling Times on the Taxonomic Composition of Permeates Microbiome

In this study, permeate represents the main external factor acting on the CW plant components; thus, we focused our attention on permeate microbial composition. The limited number of collected permeate samples did not allow us to perform statistical analyses for the three different sampling sites of the CW. For this reason, all the permeate samples collected at the same timepoint were treated as a unique group regardless of their spatial provenience to focus only on differences due to the time of sampling.

Subsequently, differential abundance analysis was performed on permeates samples during time. Differences between permeates were statistically significant, particularly in relation to the fourth and the fifth samplings, as confirmed by a PERMANOVA on the PCA to the centered log-ratio (Table S5 and Figure 4).

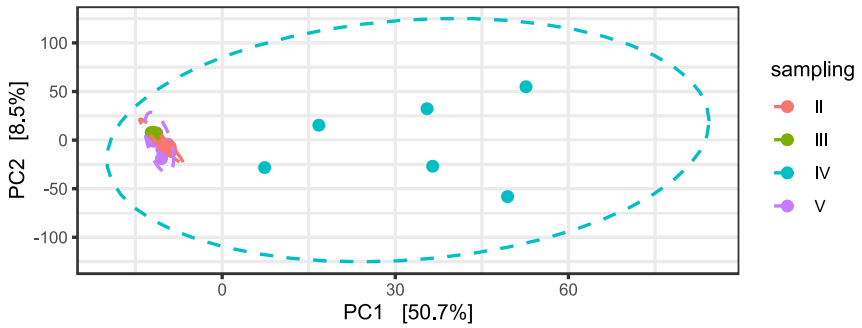


Figure 4. Principal Component Analysis (PCA) on centered log-ratio transformed counts at ASV level and variance percentage explained by each component. Ellipses show the Normal-theory confidence regions with alpha = 0.95. Permeate samples are colored in relation to the time of sampling.

Differences in permeate taxonomic composition are present in relation to 197 ASVs with statistically different abundance when considering sampling timepoint (Table S6). Most of the 197 ASVs include members of the orders Flavobacteriales, Sphingomonadales, Rhodobacterales, Oceanospirillales, and Pseudomonadales.

2.4. Prediction of Potential Functions and Metabolic Pathways of Permeates

Functional profiles of the permeate samples microbiota were then investigated. All the identified metabolic pathways present in the taxa forming the microbiota are listed in Table S7. We identified a total of 423 pathways, all distributed in all samples. However, pathway abundances showed a clustering according to sampling time, in particular for the fourth sampling (Figure 5A).

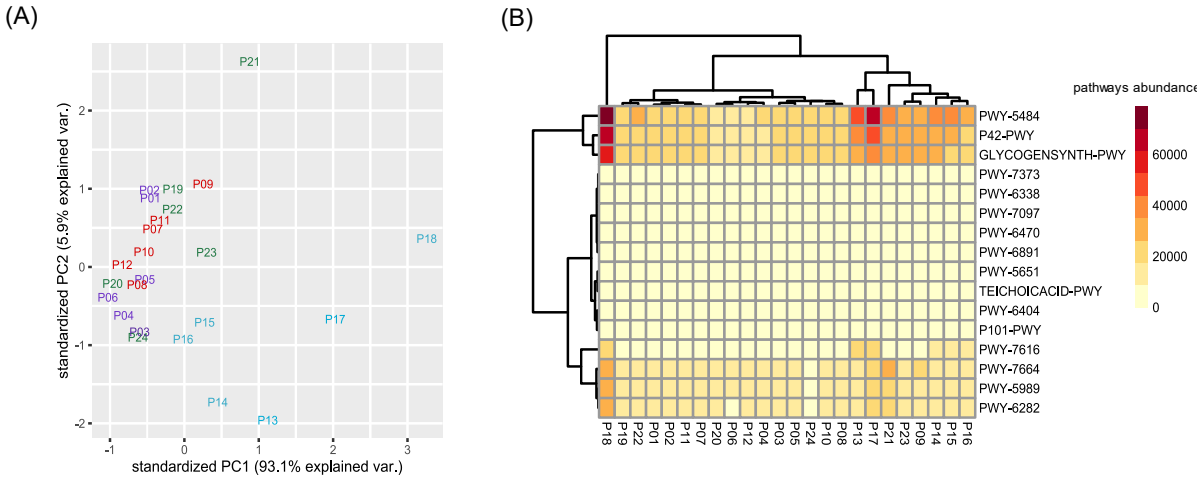


Figure 5. Clustering of pathways abundance. (**A**) Principal Component Analysis (PCA) was performed considering the predicted functional abundances of all identified pathways. In violet permeates of the second sampling, in red for the third, in blue for the fourth, and in green for the fifth. (**B**) Heatmap considering only the top three pathways that differentiate each contrast between varieties returned by SIMPER test. Codes of represented pathways refer to the MetaCyc Metabolic Pathway Database.

Using a SIMPER test, we selected the top three pathways that mostly differentiate the four sampling timepoints (in terms of amount of explained variance). For each contrast between different sampling times, the three most represented metabolic pathways were selected, for a total of sixteen different pathways (i.e., PWY-6470, PWY-7373, PWY-7664, P42-PWY, GLYCOGENSYNTH-PWY, PWY-5484, PWY-6891, PWY-6404, P101-PWY, PWY-5651, TEICHOICACID-PWY, PWY-5989, PWY-6282, PWY-6338, PWY-7097, PWY-7616),

related to the (i) metabolism of aerobic respiration, (ii) antibiotic resistance, (iii) vanillin and vanillate metabolism, and (iv) biosynthesis of stress related compounds (Table S8). The clustering of this subset of pathways agreed with the overall results of the taxonomic diversity observed for permeate samples (Figure 5A). However, a heatmap reporting the sixteen most expressed pathways did not show a particular clustering of permeate samples, with the exception of those belonging to the fourth sampling, which grouped together (Figure 5B).

3. Discussion

The aim of this work was the characterization of the whole bacterial communities associated to different compartments of the plant *P. australis*, soils, and permeates of the pilot constructed wetlands (CW) of Calice (Prato, Italy), for a total of 122 samples, to evaluate the possible effect of wastewaters on the CW bacterial communities. The composition of the total microbiota was determined through five samplings spanning from March 2017 to December 2018, with the first one performed before the activation of the plant, thus, before the entry of wastewaters.

The analysis of the diversity within each sample highlighted a higher diversity for soils (i.e., rhizosphere and bulk soil) and permeates with respect to the plant tissues (i.e., aerial part, stem, and root). Bacterial communities' compositions were similar for the different plant compartments, while differences were observed for soils and permeates, which were highly different from each other and from plant tissues samples, showing a deep dependence on the sampling time.

This result is a quite interesting aspect. Indeed, permeates continuously flow through CW, staying for relatively short times inside the tanks (in a time range spanning from few hours to few days, depending on the season, weather, etc.). On the contrary, *P. australis* plant, its associated soil, and the surrounding soil filling the tanks of CW, remain stable in the CW and are subjected to a slower renewal. It is unlikely that permeate could shape CW total microbiota completely; nonetheless, it is very probable that permeates somehow affect CW bacterial communities. The fact that soils and permeates bacterial communities resulted different from each other, despite their strict physical interaction, can be explained based on the continuous influx of different wastewaters, whose differences can be due to seasonal changes and/or origin, potentially leading even to remarkable variations. However, the effect of permeates on environmental microbiota can be highlighted by the similarity in bacterial diversity among these samples. On the other hand, *P. australis* bacterial community seems to remain stable among the different plant tissues, with minor variations along time. This is also true in the case of *P. australis* roots, which are in close contact with the environment. This aspect allows the hypothesis that *P. australis* is able to select and hold a specific microbiota, as already proposed in the case of other plants such as medicinal ones [15]. Indeed, it was previously reported that, as in the case of wastewater treatment plants, interactions between bacteria exert the main effect in determining the composition of the bacterial community in comparison to environmental properties [16,17]. Moreover, maintenance of a stable balance between host and associated microorganisms could improve plant health through prevention of dysbiosis and pathogen colonization [18]. All these considerations are particularly fitting in the case of *P. australis* growing in CW: in fact, presence of a stable microbiota could also be related to its ability to thrive in a harsh environment, characterized by the presence of diverse pollutants of various origins. From a biotechnological and applicative viewpoint, stability of the microbiota associated to the plant tissues could be exploited to shape and improve phytoremediative performances of CW. Indeed, *P. australis* could be inoculated with selected beneficial bacterial strains before its implant in CW, possibly leading to a selective advantage for those specimens having this controlled microbial flora. On the other hand, regular checks of the microbial community might be useful to predict an upcoming decrease in remediative performance due, for instance, to dysbiosis events. However, new experiments are certainly needed to confirm these hypotheses on the field, and comparisons between the bacterial communities

of *P. australis* from different CW worldwide could represent an untapped source of valuable information to improve phytodepuration efficiency.

No evident shift of the microbial communities seems associated to seasonal changes and/or meteorological factors, reinforcing the hypothesis that the permeate acts as an exogenous factor influencing changes mainly in soil microbial communities without however replicating its own composition. Under the light of possible functional differences among the bacterial communities present in the permeate, PICRUSt2 results mirrored, in part, the taxonomic profiling of permeate samples. Pathways related to antibiotic resistance were found in relation with the taxonomy of the bacterial communities associated to the permeates. This is in line with the provenience of the wastes treated in the G.I.D.A. SpA plant, designed for the tertiary treatment of landfill leachate (LFL), thus presenting material from anthropogenic provenience and potentially rich in antimicrobial compounds [19]. Natural vanillin production, evidenced by the presence of the PWY-6338 and PWY-7097 pathways, represents an important perspective to exploit the potential of bacterial communities associated to waste materials. The exploitation of biological synthetic pathways, such as those involving microorganisms, rather than relying on petroleum derivatives as major raw materials, can result in an environmentally friendly process [20].

In particular, permeates harvested during the fourth timepoint grouped separately from the others, similar to the pattern shown in the PCA analysis of ASVs (Figure 4). Interestingly, the metabolic pathways, which mostly differentiated the permeate collected during the fourth sampling from the others, were mainly related to an increased aerobic respiration and fatty acid biosynthesis (PWY-7664, PWY-5989, PWY-6282). These permeate bacterial communities contained a higher amount of Firmicutes, a phylum including several aerobic bacterial groups thriving in soil, such as *Bacillus*. Fatty acid biosynthesis has already been described in numerous studies on CW [21–23]. Hydrolysis, acidification, and solubilization of complex organic compounds are extremely influenced by environmental factors such as temperature and hydraulic retention time [24,25]. For this reason, it is possible to observe fluctuation of this pathway, since the G.I.D.A. SpA CW plant is subjected to this kind of fluctuations. Finally, pathways PWY-7616 and P42-PWY are linked to the methanol metabolism. Indeed, enzymes of an incomplete reductive TCA cycle have been experimentally demonstrated as present in the autotrophic methanogen *Methanococcus maripaludis* [26].

In conclusion, this work allowed a further characterization of the bacterial community associated to *P. australis* and its surrounding environment exploited for phytodepuration in CW. Previous results obtained on the culturable component of *P. australis* root microbiota showed that resistance against antibiotics and metal(loid)s increased in relation to the time of exposure to wastewater. Here, we observed the main changes regarding environmental samples, and that plants host a more stable bacterial community. Since analysis of the functional profiles of permeate sample microbiota revealed the presence of pathways related to antibiotic resistance, it cannot, a priori, be excluded the occurrence of horizontal gene transfer events between wastewater- and plant-associated bacteria.

4. Materials and Methods

4.1. Site Description and Samplings

Samples of *P. australis*, soils, and permeates were collected from the CW within the wastewater treatment plant in Calice (Tuscany, Italy). The site and CW were previously described [13,14]. Briefly, the CW plant consists of two independent lines, namely, "Lane A" and "Lane B", each representing a two-stage subsurface flow system (SFS), with a vertical system (SFS-v) followed by a horizontal one (SFS-h). Samples were collected during five sampling campaigns spanning 22 months (i.e., first—March 2017; second—July 2017; third—November 2017; fourth—June 2018; fifth—December 2018), with the first one performed before the activation of the CW (i.e., no permeates and LFLs were present yet). Samples were harvested using sterile plastic bags and transferred immediately to the laboratory for DNA extraction. Three plant and soil samples collected from each SFS were

pooled before proceeding with DNA extraction. Three kinds of permeates were sampled, corresponding to three different points of the CW: entrance of the SFS-v, exit from the SFS-v (i.e., entrance of the SFS-h), and exit from the SFS-h. Each sample was collected and analyzed in duplicate. Except for permeates, all samples were name-coded based on (i) the SFS of origin (i.e., "V" for SFS-v and "H" for SFS-h), (ii) the replicate (1–2, 3–4, 5–6, 7–8, and 9–10 for the first, second, third, fourth, and fifth sampling, respectively), and (iii) the kind of sample (i.e., "A", "S", "R", "RS", and "BS" for aerial part, stem, root, rhizospheric soil, and bulk soil, respectively) (Table 1). Permeates follow a progressive nomenclature based on (i) the sampling point and (ii) timing of sampling (Table 1).

Table 1. Description of the samples collected in this work, and associated identification codes.

SFS	Sample	Sampling				
		I	II	III	IV	V
SFS-v Entrance	Permeate		P1 P2	P7 P8	P13 P14	P19 P20
SFS-v	Aerial part	V01A V02A	V03A V04A	V05A V06A	V07A V08A	V09A V10A
	Stem	V01S V02S	V03S V04S	V05S V06S	V07S V08S	V09S V10S
	Root	V01R V02R	V03R V04R	V05R V06R	V07R V08R	V09R V10R
	Rhizosphere	V01RS V02RS	V03RS V04RS	V05RS V06RS	V07RS V08RS	V09RS V10RS
	Bulk soil	V01BS V02BS	V03BS V04BS	V05BS V06BS	V07BS V08BS	V09BS V10BS
SFS-v Exit SFS-h Entrance	Permeate		P3 P4	P9 P10	P15 P16	P21 P22
SFS-h	Aerial part	H01A H02A	H03A H04A	H05A H06A	H07A H08A	H09A H10A
	Stem	H01S H02S	H03S H04S	H05S H06S	H07S H08S	H09S H10S
	Root	H01R H02R	H03R H04R	H05R H06R	H07R H08R	H09R H10R
	Rhizosphere	H01RS H02RS	H03RS H04RS	H05RS H06RS	H07RS H08RS	H09RS H10RS
	Bulk soil	H01BS H02BS	H03BS H04BS	H05BS H06BS	H07BS H08BS	H09BS H10BS
SFS-h Exit	Permeate		P5 P6	P11 P12	P17 P18	P23 P24

4.2. eDNA Extraction and 16S rRNA Gene Amplicon Sequencing

Environmental DNA (eDNA) extraction was performed using the PowerLyzer PowerSoil DNA Isolation Kit (MO BIO Laboratories), according to the protocol provided by the manufacturer, under sterile conditions. Quality of DNA was assessed by 0.8% w/v agarose gel electrophoresis. Library construction and DNA sequencing of V3-V4 region of 16S rDNA were performed by IGA Technology Services (Udine, Italy). The MiSeq Illumina platform was used for DNA sequencing in the 2 × 250 paired-end format. All the sequenced reads were submitted to Sequence Read Archive (SRA) database under the accession number PRJNA885383.

4.3. Bioinformatic and Statistical Analyses

Illumina reads were trimmed using 'Cutadapt' (version 3.7) [27]. Paired-end sequences were clustered into Amplicon Sequence Variants (ASVs) following the default settings of the DADA2 pipeline (version 1.16) reported at https://benjjneb.github.io/dada2/tutorial.html

(accessed on 12 April 2022) [28], relying on the R environment (version 4.0.5) [29]. After the quality check, reads were filtered using the DADA2 'filterAndTrim' function, setting the maximum error rate to 2 and the fixed lengths to 250 and 220 bp for forward and reverse reads, respectively. Trimmed sequences were the input for error rate estimation using 'learnErrors' function with default parameters. Following denoising and inferring variants, sequences were merged [30]. After removing chimeras, taxonomic annotation was performed using the Silva database version 132 [31]. The obtained differential abundance tables were analyzed with the 'phyloseq' R package (version 1.34.0) [32]. ASV assigned to mitochondria and chloroplasts and those not assigned to Bacteria were removed.

Statistical tests on microbial communities were performed in R environment (version 4.0.5). Rarefaction curves were obtained using the function 'ggrare()' within 'ranacapa' R package (version 0.1.0) [33]. For alpha diversity analysis, the different indices were calculated and plotted using the function 'estimate_richness()' within 'phyloseq' R package (version 1.34.0) [34]. A Wilcoxon test for multiple comparison of averages was performed on alpha diversity indices (Shannon and Simpson) using the 'pairwise.wilcox.test()' function within the 'stats' R package (version 4.0.5). Good's coverage was calculated through the R functions 'goods()' within the 'QsRutils' R package (version 0.1.5) [35].

Beta-diversity was analyzed using the 'adonis()' function within the 'vegan' R package (version 2.5-7) [36]. Distances were calculated using the 'ordinate()' function within the 'phyloseq' R package (version 1.34.0) with 'bray' as distance and 'NMDS' as method settings [32]. Plots were built using the 'plot_ordination()' function within the 'phyloseq' R package (version 1.34.0). To assess statistical significance, the function 'pairwise.adonis()' within the 'pairwiseAdonis' R package (version 0.0.1) was used. Using the 'ggplot()' function within the 'ggplot2' R package (version 3.3.5) [37], plots were built. Differential abundance analysis was performed using the R package DeSeq2 v1.30.1 to identify the ASVs/taxa differentially expressed in the samples. The functions 'phyloseq_to_deseq2()', 'DESeq()' and 'results()' were used [38].

4.4. Prediction of Functional Abundances

In order to investigate and predict the functional potential for each determined sequence in microbial communities, the software PICRUSt2 (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States, https://github.com/picrust/picrust2, accessed on 12 April 2022) was used. The metabolic pathway abundances and functional gene profiling were obtained by comparison with MetaCyc and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases [39]. The most relevant pathways for SFS differentiation were ranked running a SIMPER test using the function 'simper()' with the 'vegan' R package (version 2.5.7). For each contrast between time of sampling, the top three differentially expressed pathways were selected and analyses were carried out using the functions 'prcomp()' to perform a Principal Component Analysis (PCA) and 'heatmap()' in 'stats' R package (version 4.0.5) to graphically represent results.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/xxx/s1, Dataset S1: total detected ASVs; Figure S1: rarefaction curves; Table S1: Good's coverage; Table S2: alpha diversity indices; Table S3: ANOVA on the alpha diversity measures; Table S4: PERMANOVA on the beta diversity measures; Table S5: PERMANOVA analysis of clr transformed permeate samples, Table S6: differentially abundant ASVs in permeate samples; Table S7: identified metabolic pathways; Table S8: pathways differentiating each contrast between varieties.

Author Contributions: Conceptualization, E.C., D.F., and R.F.; methodology, C.F., C.B., and E.C.; software, L.C. and M.S.; validation, L.C., M.S., and S.D.D.; formal analysis, L.C., S.D.D., and A.V.; investigation, L.C., M.S., S.D.D., C.F., and A.V.; data curation, L.C. and S.D.D.; writing—original draft preparation, L.C., S.D.D., C.F., and A.V.; writing—review and editing, L.C., M.S., S.D.D., C.F., C.B., E.C., D.F., R.F., and A.V.; visualization, L.C. and S.D.D.; supervision, D.F., R.F., and A.V.; project administration, R.F.; funding acquisition, D.F. and R.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All sequences have been submitted online. Metagenomic sequences were deposited in the NCBI Sequence Read Archive (SRA) under the accession PRJNA885383. All the scripts used for the analyses are available at https://github.com/lisacangioli/Phragmites-australis.git (accessed on 29 September 2022).

Acknowledgments: We are grateful to G.I.D.A. SpA (Gestione Impianti di Depurazione Acque, Prato, Italy), who supported this work. We would also like to thank Adriano Buzzati-Traverso Foundation (Pavia, Italy), who supported this work with a research fellowship grant to S.D.D.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Riva, V.; Riva, F.; Vergani, L.; Crotti, E.; Borin, S.; Mapelli, F. Microbial assisted phytodepuration for water reclamation: Environmental benefits and threats. *Chemosphere* **2020**, *241*, 124843. [CrossRef]
2. Vymazal, J. Constructed wetlands for wastewater treatment: Five decades of experience. *Environ. Sci. Technol.* **2011**, *45*, 61–69. [CrossRef]
3. Vassallo, A.; Kett, S.; Purchase, D.; Marvasi, M. Antibiotic-Resistant Genes and Bacteria as Evolving Contaminants of Emerging Concerns (e-CEC): Is It Time to Include Evolution in Risk Assessment? *Antibiotics* **2021**, *10*, 1066. [CrossRef] [PubMed]
4. Wu, H.; Zhang, J.; Ngo, H.H.; Guo, W.; Hu, Z.; Liang, S.; Fan, J.; Liu, H. A review on the sustainability of constructed wetlands for wastewater treatment: Design and operation. *Bioresour. Technol.* **2015**, *175*, 594–601. [CrossRef]
5. Kataki, S.; Chatterjee, S.; Vairale, M.G.; Dwivedi, S.K.; Gupta, D.K. Constructed wetland, an eco-technology for wastewater treatment: A review on types of wastewater treated and components of the technology (macrophyte, biolfilm and substrate). *J. Environ. Manage.* **2021**, *283*, 111986. [CrossRef] [PubMed]
6. Kurniawan, S.B.; Ahmad, A.; Said, N.S.M.; Imron, M.F.; Abdullah, S.R.S.; Othman, A.R.; Purwanti, I.F.; Hasan, H.A. Macrophytes as wastewater treatment agents: Nutrient uptake and potential of produced biomass utilization toward circular economy initiatives. *Sci. Total Environ.* **2021**, *790*, 148219. [CrossRef] [PubMed]
7. Bianchi, E.; Coppi, A.; Nucci, S.; Antal, A.; Berardi, C.; Coppini, E.; Fibbi, D.; Del Bubba, M.; Gonnelli, C.; Colzi, I. Closing the loop in a constructed wetland for the improvement of metal removal: The use of Phragmites australis biomass harvested from the system as biosorbent. *Environ. Sci. Pollut. Res.* **2021**, *28*, 11444–11453. [CrossRef] [PubMed]
8. Kumar Yadav, K.; Gupta, N.; Kumar, A.; Reece, L.M.; Singh, N.; Rezania, S.; Ahmad Khan, S. Mechanistic understanding and holistic approach of phytoremediation: A review on application and future prospects. *Ecol. Eng.* **2018**, *120*, 274–298. [CrossRef]
9. Wang, J.; Long, Y.; Yu, G.; Wang, G.; Zhou, Z.; Li, P.; Zhang, Y.; Yang, K.; Wang, S. A Review on Microorganisms in Constructed Wetlands for Typical Pollutant Removal: Species, Function, and Diversity. *Front. Microbiol.* **2022**, *13*, 925. [CrossRef]
10. Yan, A.; Wang, Y.; Tan, S.N.; Mohd Yusof, M.L.; Ghosh, S.; Chen, Z. Phytoremediation: A Promising Approach for Revegetation of Heavy Metal-Polluted Land. *Front. Plant Sci.* **2020**, *11*, 359. [CrossRef]
11. Vymazal, J. Emergent plants used in free water surface constructed wetlands: A review. *Ecol. Eng.* **2013**, *61*, 582–592. [CrossRef]
12. Rezania, S.; Park, J.; Rupani, P.F.; Darajeh, N.; Xu, X.; Shahrokhishahraki, R. Phytoremediation potential and control of Phragmites australis as a green phytomass: An overview. *Environ. Sci. Pollut. Res.* **2019**, *26*, 7428–7441. [CrossRef] [PubMed]
13. Coppini, E.; Palli, L.; Antal, A.; Del Bubba, M.; Miceli, E.; Fani, R.; Fibbi, D. Design and start-up of a constructed wetland as tertiary treatment for landfill leachates. *Water Sci. Technol.* **2019**, *79*, 145–155. [CrossRef]
14. Vassallo, A.; Miceli, E.; Fagorzi, C.; Castronovo, L.M.; Del Duca, S.; Chioccioli, S.; Venditto, S.; Coppini, E.; Fibbi, D.; Fani, R. Temporal Evolution of Bacterial Endophytes Associated to the Roots of Phragmites australis Exploited in Phytodepuration of Wastewater. *Front. Microbiol.* **2020**, *11*, 1652. [CrossRef] [PubMed]
15. Castronovo, L.M.; Vassallo, A.; Mengoni, A.; Miceli, E.; Bogani, P.; Firenzuoli, F.; Fani, R.; Maggini, V. Medicinal Plants and Their Bacterial Microbiota: A Review on Antimicrobial Compounds Production for Plant and Human Health. *Pathogens* **2021**, *10*, 106. [CrossRef]
16. Cydzik-Kwiatkowska, A.; Zielińska, M. Bacterial communities in full-scale wastewater treatment systems. *World J. Microbiol. Biotechnol.* **2016**, *32*, 66. [CrossRef]
17. Ju, F.; Zhang, T. Bacterial assembly and temporal dynamics in activated sludge of a full-scale municipal wastewater treatment plant. *ISME J.* **2015**, *9*, 683–695. [CrossRef]
18. Getzke, F.; Thiergart, T.; Hacquard, S. Contribution of bacterial-fungal balance to plant and animal health. *Curr. Opin. Microbiol.* **2019**, *49*, 66–72. [CrossRef]
19. Vassallo, A.; Kett, S.; Purchase, D.; Marvasi, M. The Bacterial Urban Resistome: Recent Advances. *Antibiotics* **2022**, *11*, 512. [CrossRef]

20. Zhao, X.; Chen, J.; Meng, X.; Li, L.; Zhou, X.; Li, J.; Bai, S. Environmental profile of natural biological vanillin production via life cycle assessment. *J. Clean. Prod.* **2021**, *308*, 127399. [CrossRef]
21. Huang, X.; Zhao, J.; Xu, Q.; Li, X.; Wang, D.; Yang, Q.; Liu, Y.; Tao, Z. Enhanced volatile fatty acids production from waste activated sludge anaerobic fermentation by adding tofu residue. *Bioresour. Technol.* **2019**, *274*, 430–438. [CrossRef] [PubMed]
22. Greses, S.; Tomás-Pejó, E.; Gónzalez-Fernández, C. Agroindustrial waste as a resource for volatile fatty acids production via anaerobic fermentation. *Bioresour. Technol.* **2020**, *297*, 122486. [CrossRef]
23. Owusu-Agyeman, I.; Plaza, E.; Cetecioglu, Z. Production of volatile fatty acids through co-digestion of sewage sludge and external organic waste: Effect of substrate proportions and long-term operation. *Waste Manag.* **2020**, *112*, 30–39. [CrossRef]
24. Fang, Q.; Ji, S.; Huang, D.; Huang, Z.; Huang, Z.; Zeng, Y.; Liu, Y. Impact of Alkaline Pretreatment to Enhance Volatile Fatty Acids (VFAs) Production from Rice Husk. *Biochem. Res. Int.* **2019**, *2019*, 8489747. [CrossRef]
25. Cheah, Y.K.; Vidal-Antich, C.; Dosta, J.; Mata-Álvarez, J. Volatile fatty acid production from mesophilic acidogenic fermentation of organic fraction of municipal solid waste and food waste under acidic and alkaline pH. *Environ. Sci. Pollut. Res.* **2019**, *26*, 35509–35522. [CrossRef]
26. Shieh, J.S.; Whitman, W.B. Pathway of acetate assimilation in autotrophic and heterotrophic methanococci. *J. Bacteriol.* **1987**, *169*, 5327–5329. [CrossRef] [PubMed]
27. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **2011**, *17*, 10–12. [CrossRef]
28. Norman Network. Available online: https://www.norman-network.net/ (accessed on 26 May 2021).
29. R: The R Project for Statistical Computing. Available online: https://www.r-project.org/ (accessed on 27 September 2022).
30. Callahan, B.J.; McMurdie, P.J.; Holmes, S.P. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* **2017**, *11*, 2639–2643. [CrossRef] [PubMed]
31. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, D590–D596. [CrossRef]
32. McMurdie, P.J.; Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **2013**, *8*, e61217. [CrossRef]
33. Kandlikar, G.S.; Gold, Z.J.; Cowen, M.C.; Meyer, R.S.; Freise, A.C.; Kraft, N.J.B.; Moberg-Parker, J.; Sprague, J.; Kushner, D.J.; Curd, E.E. Ranacapa: An R package and shiny web app to explore environmental DNA data with exploratory statistics and interactive visualizations [version 1; referees: 1 approved, 2 approved with reservations]. *F1000Research* **2018**, *7*, 1734. [CrossRef] [PubMed]
34. McMurdie, P.J.; Holmes, S. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLOS Comput. Biol.* **2014**, *10*, e1003531. [CrossRef]
35. Quensen, J. QsRutils: R Functions Useful for Community Ecology. 2020. Available online: https://github.com/jfq3/QsRutils (accessed on 13 April 2022).
36. Dixon, P. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* **2003**, *14*, 927–930. [CrossRef]
37. Wickham, H. *ggplot2*; Springer: Cham, Switzerland, 2016. [CrossRef]
38. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef] [PubMed]
39. Douglas, G.M.; Maffei, V.J.; Zaneveld, J.R.; Yurgel, S.N.; Brown, J.R.; Taylor, C.M.; Huttenhower, C.; Langille, M.G.I. PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* **2020**, *38*, 685–688. [CrossRef]

International Journal of
Molecular Sciences

MDPI

Article

Lung and Gut Microbiota Changes Associated with *Pseudomonas aeruginosa* Infection in Mouse Models of Cystic Fibrosis

Giovanni Bacci [1], Alice Rossi [2], Federica Armanini [3], Lisa Cangioli [1], Ida De Fino [2], Nicola Segata [3], Alessio Mengoni [1], Alessandra Bragonzi [2] and Annamaria Bevivino [4,*]

1 Department of Biology, University of Florence, Sesto Fiorentino, 50019 Florence, Italy; giovanni.bacci@unifi.it (G.B.); lisa.cangioli@unifi.it (L.C.); alessio.mengoni@unifi.it (A.M.)
2 Infections and Cystic Fibrosis Unit, Division of Immunology, Transplantation and Infectious Diseases, IRCCS San Raffaele Scientific Institute, 20132 Milan, Italy; rossi1.alice@hsr.it (A.R.); defino.ida@hsr.it (I.D.F.); bragonzi.alessandra@hsr.it (A.B.)
3 Department CIBIO, University of Trento, 38122 Trento, Italy; federica.armanini@unitn.it (F.A.); nicola.segata@unitn.it (N.S.)
4 Department for Sustainability, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, ENEA Casaccia Research Center, 00123 Rome, Italy
* Correspondence: annamaria.bevivino@enea.it; Tel.: +39-0630-483-868

Abstract: Cystic fibrosis (CF) disease leads to altered lung and gut microbiomes compared to healthy subjects. The magnitude of this dysbiosis is influenced by organ-specific microenvironmental conditions at different stages of the disease. However, how this gut-lung dysbiosis is influenced by *Pseudomonas aeruginosa* chronic infection is unclear. To test the relationship between CFTR dysfunction and gut-lung microbiome under chronic infection, we established a model of *P. aeruginosa* infection in wild-type (WT) and gut-corrected CF mice. Using 16S ribosomal RNA gene, we compared lung, stool, and gut microbiota of C57Bl/6 *Cftr* ^{tm1UNC}TgN(FABP*CFTR*) or WT mice at the naïve state or infected with *P. aeruginosa*. *P. aeruginosa* infection influences murine health significantly changing body weight both in CF and WT mice. Both stool and gut microbiota revealed significantly higher values of alpha diversity in WT mice than in CF mice, while lung microbiota showed similar values. Infection with *P. aeruginosa* did not changed the diversity of the stool and gut microbiota, while a drop of diversity of the lung microbiota was observed compared to non-infected mice. However, the taxonomic composition of gut microbiota was shown to be influenced by *P. aeruginosa* infection in CF mice but not in WT mice. This finding indicates that *P. aeruginosa* chronic infection has a major impact on microbiota diversity and composition in the lung. In the gut, CFTR genotype and *P. aeruginosa* infection affected the overall diversity and taxonomic microbiota composition, respectively. Overall, our results suggest a cross-talk between lung and gut microbiota in relation to *P. aeruginosa* chronic infection and CFTR mutation.

Keywords: cystic fibrosis; lung; gut; microbiome; gut-lung axis; animal models; CFTR mice; *Pseudomonas aeruginosa*

check for updates

Citation: Bacci, G.; Rossi, A.; Armanini, F.; Cangioli, L.; De Fino, I.; Segata, N.; Mengoni, A.; Bragonzi, A.; Bevivino, A. Lung and Gut Microbiota Changes Associated with *Pseudomonas aeruginosa* Infection in Mouse Models of Cystic Fibrosis. *Int. J. Mol. Sci.* **2021**, *22*, 12169. https://doi.org/10.3390/ijms222212169

Academic Editor: Svein-Ole Mikalsen

Received: 30 September 2021
Accepted: 5 November 2021
Published: 10 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Cystic fibrosis (CF) is the most common autosomal recessive disease caused by mutations in the Cystic Fibrosis Transmembrane Conductance Regulator (*CFTR*) gene [1]. Dysfunction of *CFTR* affects several organs, but primarily complications in the lung are the causes of morbidity and mortality [2]. Absence of CFTR-mediated Cl-secretion has been functionally linked to airway surface dehydration, which leads to the accumulation of concentrated mucus, airway obstruction, inflammation and infection, bronchiectasis, and ultimately death. Gastrointestinal (GI) manifestations are the most important non-pulmonary manifestations of CF and commonly complicate the care of CF patients. GI

tract manifestations of CF are related to mucous inspissation and dysmotility and include meconium ileus, constipation, distal intestinal obstruction syndrome, gastroesophageal reflux disease, and small bowel bacterial overgrowth. In CF, the central role of lung and gut microbiomes in the pathophysiology has been supported by several studies [3]. CF patients show altered airway microbiome, correlated to lung-specific microenvironmental conditions (e.g., viscous mucus, hyperinflammation, and infection) and altered fecal microbiomes, correlated with gastrointestinal inflammation and nutrient malabsorption [3,4]. Emerging findings show a link between gut microbiome and progression of lung disease in CF [5,6] similar to other lung diseases such as asthma, COPD, allergy, and lung cancer [7–9]. The altered composition of the gut microbiome has been shown to profoundly impact the lung inflammatory response in the context of pulmonary diseases via the vital cross-talk between the lungs and gut microbiota [10–12]. Evidence that CFTR mutations have an effect on gut microbiota are increasing in humans and in CF mice [13,14], although it is unclear how this gut-lung dysbiosis is influenced by *Pseudomonas aeruginosa* chronic infection.

Animal models of CF are of critical importance as they contribute to our understanding of the mechanisms associated with CF disease progression and disease pathology [13]. Several CF mouse models have been characterized and reported [14]. In these mice, the *Cftr* gene is either completely knocked out or harbors mutations observed in the human population. Although human and murine CFTR have many properties in common and share 78% overall sequence identity at an amino acid level [15], some important differences in function, including channel-gating behavior, have been described [16]. CF mouse models demonstrate broad evidence of intestinal disease, but exhibit large variations in survival, anatomically confined CF ion transport defect, and general absence of CF-like lung disease. These features make them difficult to work with, especially if large numbers are required. Working with CF mice has become substantially easier after the generation of the gut-corrected CF-knockout mouse. A double-mutant CF mouse has been generated, based on the UNC-null mouse (*Cftr*^{tm1Unc}) that also carries a human *CFTR* transgene controlled by rat intestinal fatty acid-binding protein gene promoter (FABP) [*Cftr*^{tm1}UNCTgN(FABP*CFTR*) [17]]. FABP-mediated expression of h*CFTR* in the gut, prevents the intestinal disease observed in CF null mice. Different CF mouse models (BALB/c *Cftr* ^{tm1Unc}, C57BL/6 *Cftr* ^{tm1Unc} and FVB/N *Cftr*^{tm1Eur} mice) revealed that intestinal disease phenotypes were associated with specific constituents of the intestinal microbiome. Analysis of CF mouse intestinal microbiota highlighted bacterial overgrowth and a decrease in microbiome richness and diversity [18–20]. Although the analysis of microbiome composition in the murine lung and gut and its dependence on CFTR mutation have been explored, the influence of *P. aeruginosa* infection has still not been investigated.

Animal models of chronic *P. aeruginosa* pneumonia have been generated to mimic the advanced stage of lung pathology in humans [14,21,22]. Chronic infection in these models is usually established by including bacterial cells in immobilizing agents where they appear to grow in microcolonies under micro-aerobic/anaerobic conditions, as they do in the mucus of CF patients. Although mild inflammatory reactions were observed with sterile agar beads, pronounced changes within the bronchi and lung parenchyma were associated with infection. Lung pathology associated with chronic experimental infection is characterized by bronchopneumonia, bronchiectasis, epithelial metaplasia, fibrosis, and alveolar exudates with inflammatory cells.

In this study, *Cftr* ^{tm1UNC}TgN(FABP*CFTR*) (named KO mice) and their wild-type (WT) congenic mice were used to investigate the influence of chronic *P. aeruginosa* infection on gut and lung microbiota. Given the potential cross-talk between *P. aeruginosa* and commensal bacteria, we addressed whether changes in respective lung and gut microbiota compositions might occur during *P. aeruginosa* colonization and whether these changes are mouse model-dependent.

2. Results

2.1. P. aeruginosa Infection Influences Murine Health Both in KO and WT Mice

KO and WT congenic mice were infected with *P. aeruginosa* RP73 strain embedded in agar beads to induce chronic infection. A group of control mice were challenged with empty beads reproducing the agar-induced inflammation and stress. Both KO and WT mice infected with *P. aeruginosa* lost significantly more body weight after 2–3 days from challenge compared to empty beads, followed by a recovery at day 7 post-infection (Figure 1A). Groups of mice challenged with *P. aeruginosa* and empty beads were significantly different in body weight changes over 6 days observation, while no significance was detected comparing KO and WT mice within the group. At day 7 post-infection, no significant differences were observed between the groups—*P. aeruginosa* infected and empty beads or KO and WT mice. At day 7 post-infection, mice were sacrificed; lung, stool, and the final straight portion of the rectum, including mucosae and feces, were aseptically excised. When *P. aeruginosa* load was evaluated in the lung, the trend of CFUs increased (not significantly, $p > 0.05$) in KO mice compared to WT mice (Figure 1B).

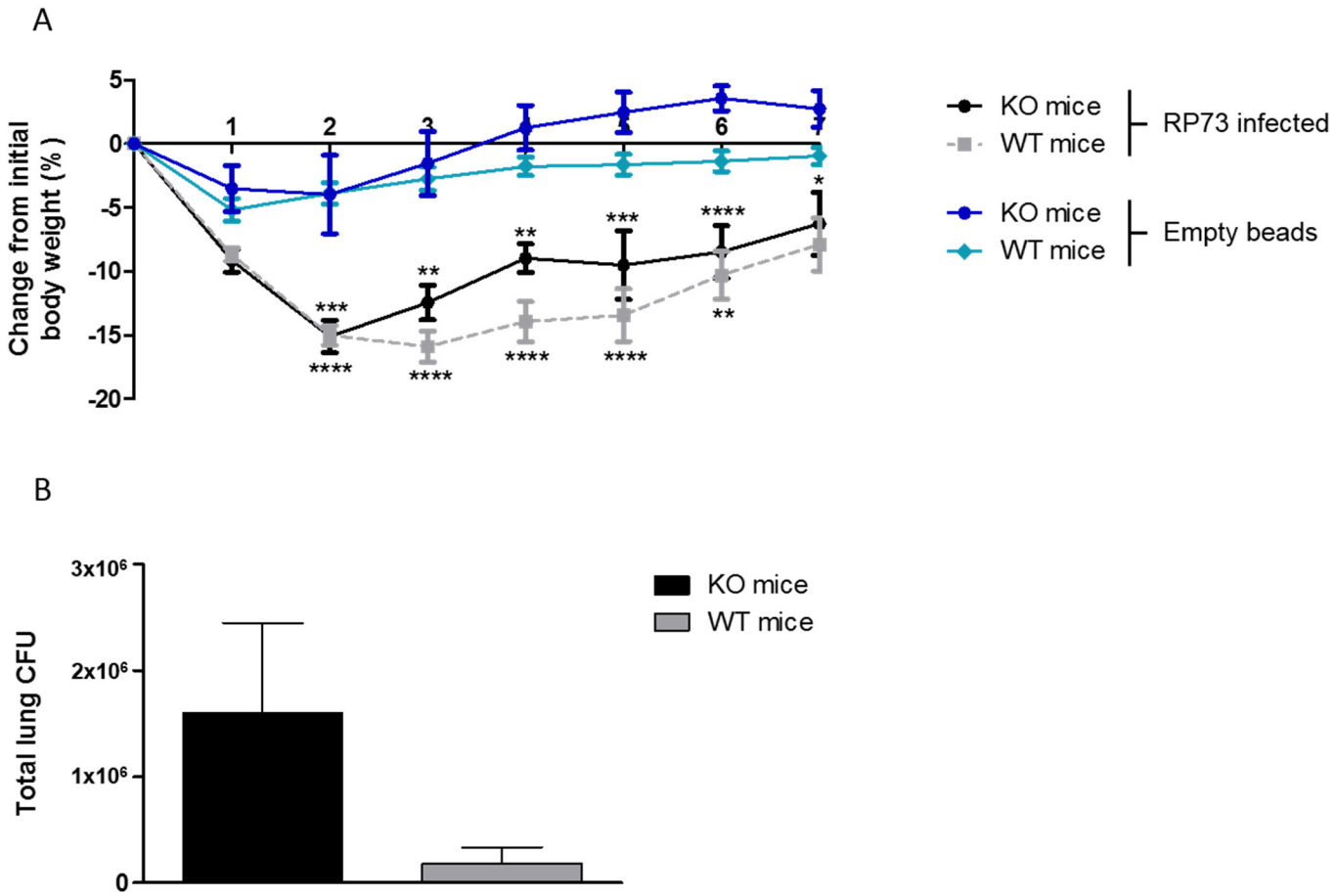


Figure 1. Body weight changes and *P. aeruginosa* load in KO and WT congenic mice. KO male and WT male mice were infected with 3–4 × 10^5 CFU of *P. aeruginosa* RP73 embedded in agar beads by intratracheal inoculation. A control group was challenged with empty beads. Mice were weighed daily, and the percentage change from the initial body weight was averaged for each group of mice (**A**). After day 7 post-inoculation, the mice were sacrificed, lungs were excised, homogenized and plated onto tryptic soy agar to determine the bacterial load (**B**). Data are presented as mean ± SEM. The data were pooled from four independent experiments (RP73 infected: KO mice $n = 13$ and WT mice $n = 17$; empty beads: KO mice $n = 7$ and WT mice $n = 8$). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. Comparisons between KO *P. aeruginosa* infected vs. KO mice challenged with empty beads and WT *P. aeruginosa* infected vs. WT mice challenged with empty beads were performed by two-way ANOVA with Bonferroni's post-test.

2.2. Microbiota Diversity Is Influenced by CFTR Genotype in the Gut and P. aeruginosa Infection in the Lung

The 16S rRNA gene amplicon sequencing reads (Table S1) were clustered into ASVs, and a total number of 10562 ASVs was retrieved (Table S2). Rarefaction curves (Figure S1) indicate a satisfactory level of sample diversity (Good's coverage, ranging from 0.99 to 1 Table S1). Alpha diversity indices in stool and gut (Table 1) showed higher values in WT mice compared to KO mice. Of note, CFTR correction of KO mice does not rescue the gut phenotype in relation to the microbiota diversity. Lung microbiota diversity was similar in KO and WT mice, mainly due to the very small number of ASVs. In KO mice, *P. aeruginosa* infection does not influence the microbiota diversity in either stool or gut compared to the mice treated with empty beads. However, lower values of diversity in the lung microbiota occurred in KO and WT mice infected with *P. aeruginosa* compared to non-infected mice (empty beads) ($p < 0.3$). Thus, the massive colonization by the *P. aeruginosa* strain has a significant influence on lung microbiota.

Table 1. Alpha diversity of mice microbiota. WT and KO mice not infected (empty beads) and infected with *P. aeruginosa* (RP73) are shown (WT mice: empty beads $n = 8$ and RP73 infected $n = 17$; KO mice: empty beads $n = 7$ and RP73 infected $n = 13$). Values (means ± standard deviation) of the Inverse Simpson index are reported.

	Gut	Stool	Lung
WT (empty beads)	37.26 ± 12.78	50.43 ± 24.48	33.45 ± 31.63
WT (RP73 infected)	41.52 ± 20.21	30.88 ± 21.48	15.02 ± 7.81
KO (empty beads)	9.81 ± 10.93	16.73 ± 10.69	40.69 ± 28.24
KO (RP73 infected)	19.12 ± 9.93	18.44 ± 5.79	15.21 ± 7.90

Clustering of samples resulted in a strong separation of lung microbiota from the intestinal (gut and stool) microbiota, while KO and CF mice were not strongly separated (Figure 2).

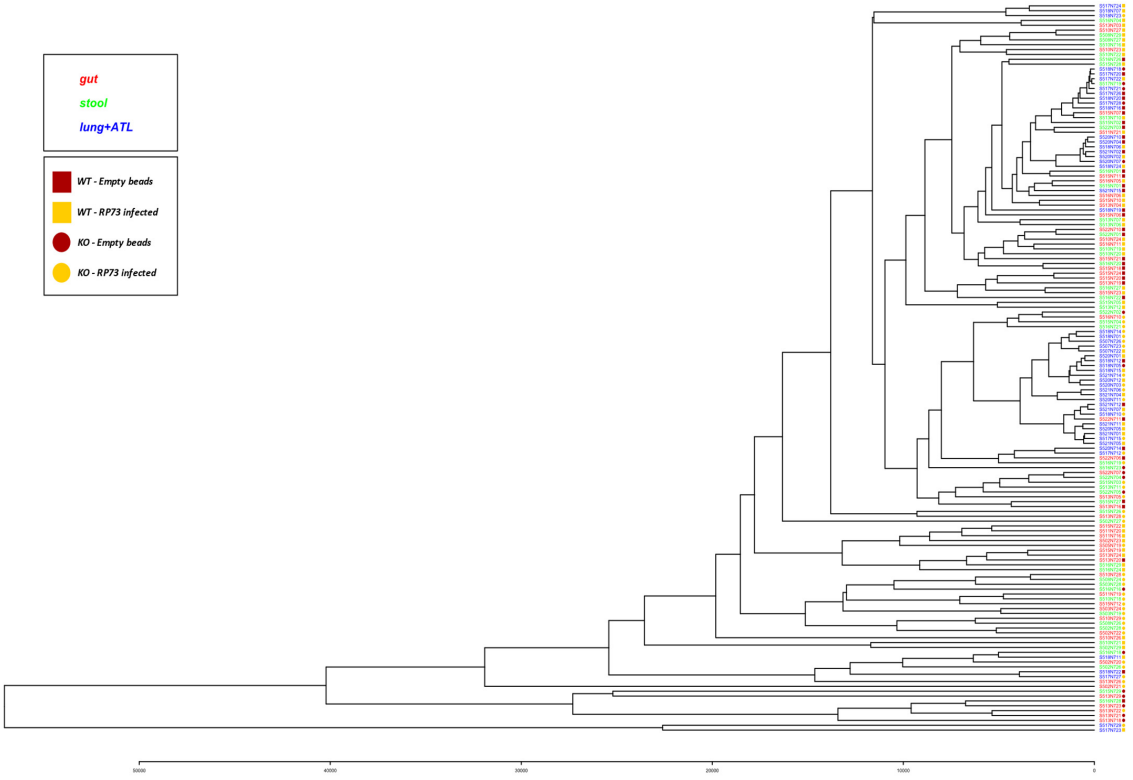


Figure 2. Clustering of mice microbiota. Codes of samples are as those reported in Table S1.

The taxonomic composition of the microbiota highlighted the division between gut district (gut and stool samples) and lung, with a peculiar grouping of taxa prevalence (Figure 3). Members of the genera *Lactobacillus*, *Akkermansia*, and the family *Lachnospiraceae* were among the most abundant taxa found. The group of greatest prevalence includes many ASVs affiliated to *Lactobacillus* and *Bacteroidales*.

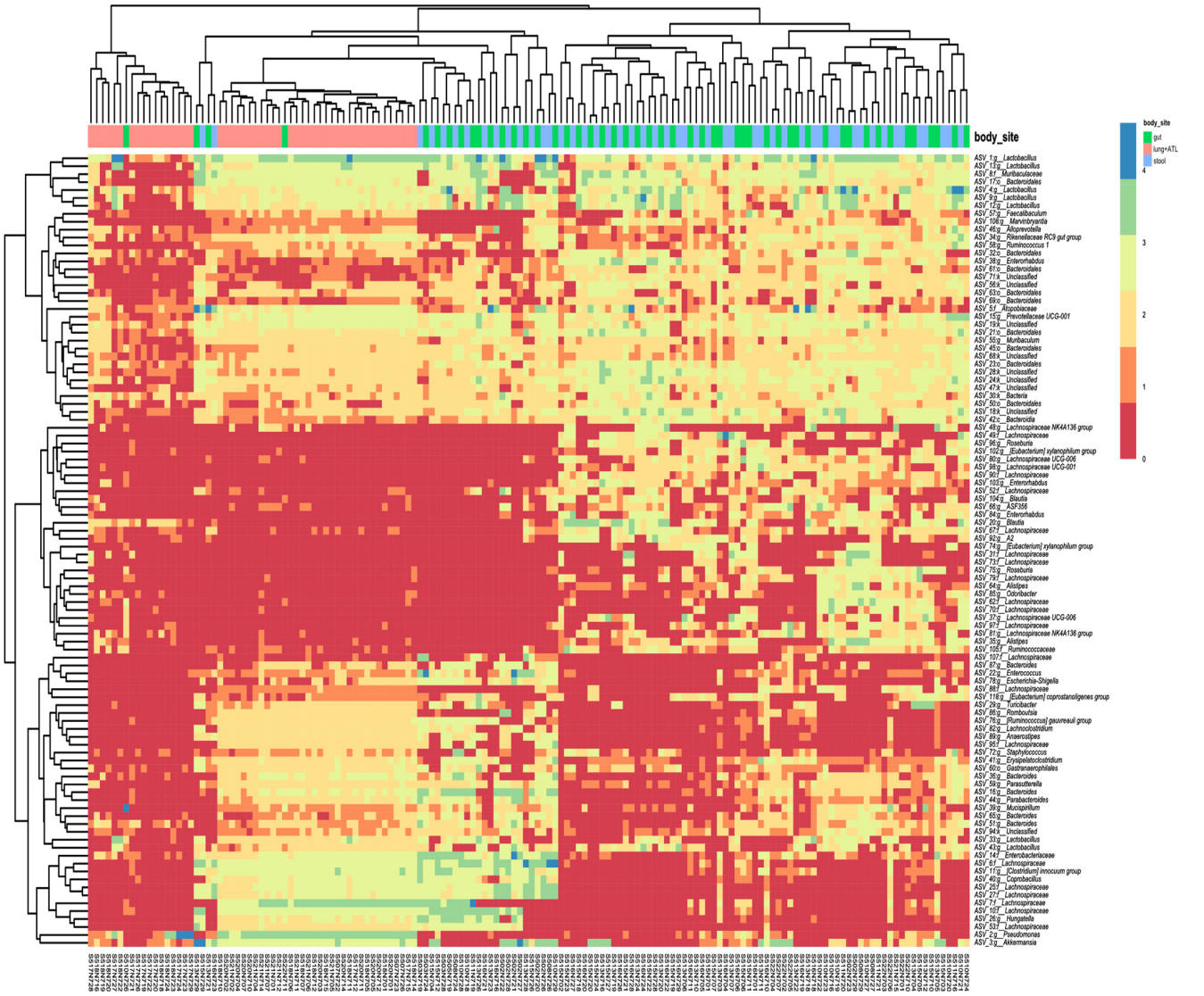


Figure 3. Taxonomic composition and clustering of samples based on the most abundant ASV (top 100). The heat-map showing the abundance of the top 100 ASVs in samples is reported.

2.3. Lung Microbiota Composition Is Strongly Influenced by *P. aeruginosa* Infection

When gut, stool, and lung microbiota composition were analysed separately, our data showed that *P. aeruginosa* infection had a more profound impact on lung microbiota than on gut and stool microbiota (Figure 4; Table 2). In fact, most of *P. aeruginosa* infected and control mice clustered separately for lung, while less sharp clustering was detected in gut and stool (Figure 2). This evidence was further confirmed by PERMANOVA (Table 2), that indicated on the overall dataset (including both KO and WT mice, Table 2a) highly significant R^2 values for *P. aeruginosa* infected compared to control mice in lung ($p < 0.001$, $R^2 = 0.099$) and more mild differences for stool ($p < 0.05$, $R^2 = 0.03$) and gut ($p < 0.05$, $R^2 = 0.04$). Interestingly, on the overall dataset (irrespective of infection) WT and KO

differed on gut ($p < 0.01$, $R^2 = 0.13$) and stool ($p < 0.01$, $R^2 = 0.17$) but not on the lung microbiota ($p > 0.8$, $R^2 = 0.01$).

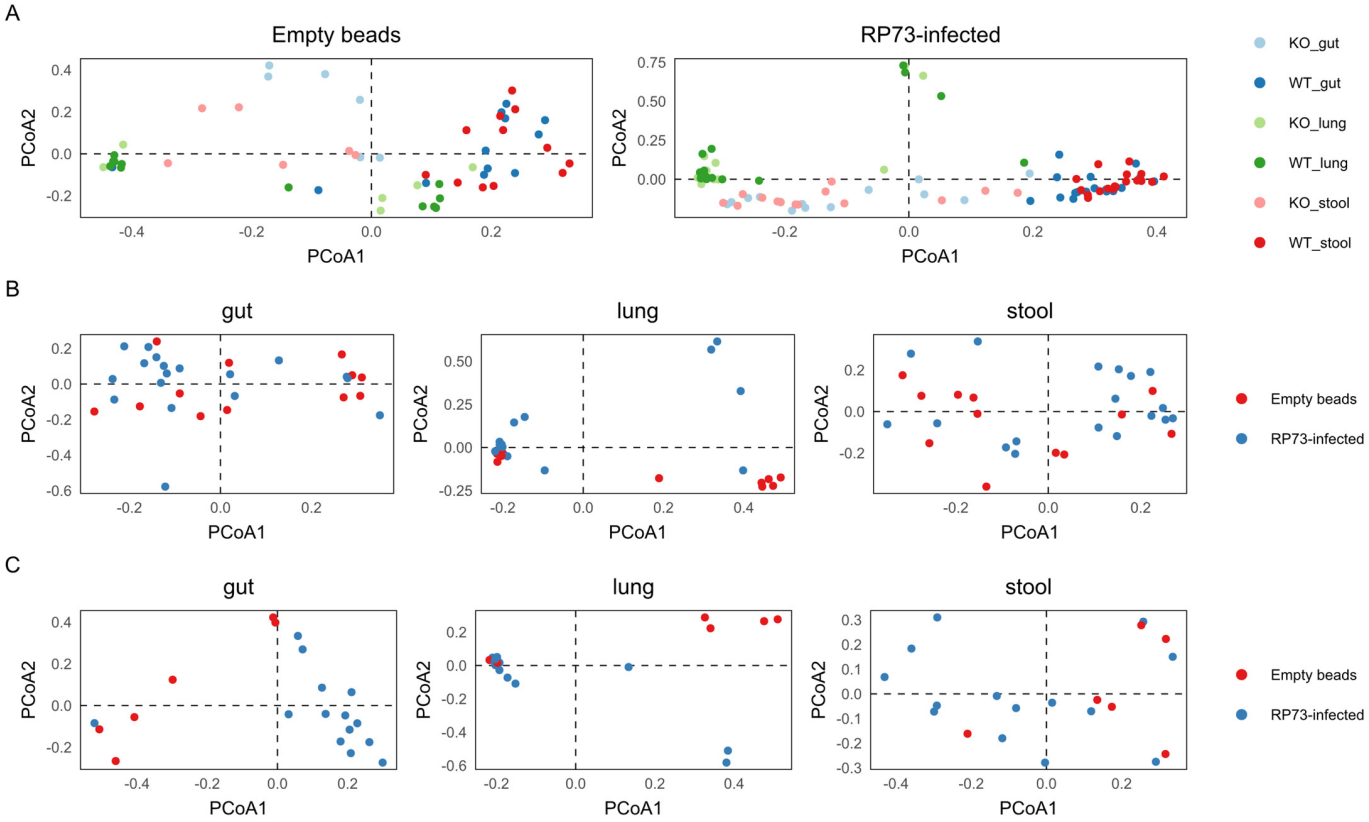


Figure 4. Principal component analysis (PCA) of mice microbiota. Pales shows the distribution of samples with respect to (**A**) infections; (**B**) WT mice and (**C**) KO mice in relation to districts. WT mice: empty beads $n = 8$ and RP73 infected $n = 17$; KO mice: empty beads $n = 7$ and RP73 infected $n = 13$.

Table 2. Infection significantly impacts on both lung and gut microbiota. Impact of body site, genotype and infection on the overall dataset (a); or restricted to RP73 infected compared to controls (b). Results of Permanova on the microbiota composition evaluating the differences among conditions and genotypes. The probability (Pr > F after F statistic calculation) and R^2 values are reported for each body site (gut, stool, lung). WT mice: empty beads $n = 8$ and RP73 infected $n = 17$; KO mice: empty beads $n = 7$ and RP73 infected $n = 13$. Asterisks indicate significant values at p-values <0.05 *, 0.01 **, 0.001 ***.

		Gut		Stool		Lung	
		Pr(>F)	R2	Pr(>F)	R2	Pr(>F)	R2
(a)	Condition (RP73 infected—control)	0.006 **	0.04305	0.026 *	0.03572	0.001 ***	0.09996
	Genotype (KO—WT)	0.001 ***	0.13225	0.001 ***	0.17754	0.894	0.01035
(b)	WT (RP73 infected—control)	0.098	0.04679	0.097	0.0461	0.023 *	0.08876
	KO (RP73 infected—control)	0.002 **	0.16172	0.100	0.0819	0.007 **	0.17221

2.4. Gut Microbiota Composition of KO Mice, but Not WT Mice, Is Influenced by P. aeruginosa Infection

Further analysis on KO and WT mice separately (Table 2b) confirmed that infection by *P. aeruginosa* impacted the gut ($p < 0.01$, $R^2 = 0.17$) and lung ($p < 0.01$, $R^2 = 0.16$) microbiota of KO mice, but only mildly the lung microbiota of WT mice ($p < 0.05$, $R^2 = 0.08$). A differential abundance analysis carried out using DeSeq2 v1.30.0 identified the ASVs/taxa differentiating the microbiota in the gut of infected mice from the CF gut control mice (adjusted p value < 0.05). We found 55 differentially represented ASVs (Table S3). Most of them belonged to the *Firmicutes* phylum.

3. Discussion

Among the many facets of CF, the microbiological status of patients is of great interest due to the recurrent, chronic microbial infection of the airways. In addition to airway microbiome alteration, CF patients show altered fecal microbiomes, correlated with gastrointestinal inflammation and nutrient malabsorption [4]. The association between lung and gut microbiome appears to be important in the pathogenesis of lung conditions as gut microbiota has been shown to influence lung inflammatory responses in CF [3,5,23]. Until now, it has been unclear whether alterations in the gut microbial species is the cause or effect of lung diseases. Investigating how gut microbiota impact other distant organs opens up new hopes for the development of therapeutic approaches to CF respiratory infections [24]. As a matter of fact, due to the complex cross-talk, the causality between lung diseases and gut microbiota is still under explored.

In this study, we investigated the lung and gut microbiome in KO and WT mice following *P. aeruginosa* infection to examine how the murine microbiome influences *P. aeruginosa* infection and its dependence on the CFTR genotype.

Several CF mouse models are available, but all the animals have severe obstructive gut disease. Differently to the human disease, the murine gut phenotype leads to high pre- and post-weaning mortality. Thus, to exclude gut pathology as a possible confounding factor on lung phenotype, we selected the gut-corrected CF mice *Cftr* ^{tm1UNC}TgN(FABP*CFTR*). We observed that, following *P. aeruginosa* infection, the bacterial loads recovered from the murine airways were higher for CF mice compared to WT mice, even if the differences were not statistically significant. The weight loss in *P. aeruginosa*-infected mice was associated with a decrease in Firmicutes phyla abundance, suggesting that *P. aeruginosa* can play a possible role in driving the biodiversity of the gut flora. These results were in agreement with other works which found that significantly reduced levels of Firmicutes correlated to weight loss in mice and humans [25,26].

Using murine models of CF infection, we compared gut and stool dysbiosis in CF mice microbiota to that of healthy WT mice with a higher microbial diversity in both gut and stool, with similar data observed for both murine models [20] and patients with CF [27]. As stated by Meeker et al. [28], CFTR dysregulation drives active selection of the gut microbiome suggesting that restoring CFTR function could favor the restoration of a healthy gut microbiota to that found in healthy, non-CF people. No difference in alpha diversity indices of the lung microbial communities was found between WT and KO mice, mainly due to the very small number of ASVs retrieved. CF mice infected with *P. aeruginosa* showed a lower diversity in lung microbiota with respect to the uninfected CF mice, most probably due to the massive colonization by *P. aeruginosa*, while no altered microbial diversity was found in the gut district. Lung microbiota composition is strongly impacted by *P. aeruginosa* infection as revealed by PCA and PERMANOVA analyses. *P. aeruginosa* infection had an effect on gut microbiota of CF mice, but not on WT mice, suggesting that the pulmonary chronic infection in CF mice may lead to altered intestinal mucosa in CF mice, irrespective of CFTR lack of expression [28]. Finding that in CF mice, gut microbiota alone (and not stool) was affected by *P. aeruginosa* raised the question of whether the mucosa-adhering microbial biofilms in the gut could be more susceptible to alteration of mucosal trophism/inflammation. We cannot exclude that, in the longer term, stool microbiota may also display the same changes. In mice, the loss of *Cftr* gene function was found to be responsible for the intestinal dysbiosis [3] and a close relationship between *Cftr* genotype and microbiome constitution was demonstrated [28]. In this study, the pulmonary chronic infection caused the alteration of intestinal mucosa in CF mice, not related to a lack of CFTR expression in the gut. We can hypothesize the existence of an "axis" from lung to gut, that is not directly related to CF defect directly, but to the inflammation status. It has been well reported that the respiratory tract has its own microbiota and lung inflammation can lead to intestinal dysbiosis [29]. Following *P. aeruginosa* infection, the differential represented by ASVs in gut microbiota, mainly belonging to Firmicutes phylum, following *P. aeruginosa* infection, underline the role of this group of bacteria with known roles in

immune modulation of the gut microbiome. A recent opinion paper [30] highlighted the importance of the mouse model for the study of gut to lung axis in CF including the study of chronic infection with agar beads. However, the exact mechanisms of microorganism selection by mice genotype are unknown. Future studies should investigate the effect of *P. aeruginosa* infection on the interaction between pathogenic bacteria and intestinal microbiota. Since the host–microbe interaction includes core microbial metabolisms and virulence factors, as well as the host's defensive countermeasures, meta'omics studies can be envisaged to fully interpret results from the *P. aeruginosa* chronic infection mouse model in a systems-biology framework [31]. Moreover, since *P. aeruginosa* strains exhibit extensive genome variation and evolution in patients' airways [32,33], experiments with different *P. aeruginosa* strains or mutant collections would be helpful to elucidate possible bacteria genotype × host genotype interactions.

4. Materials and Methods

4.1. Ethic Statement

Animal studies adhered to the Italian Ministry of Health guidelines for the use and care of experimental animals (IACUC 908).

4.2. Mouse Model and P. aeruginosa Chronic Infection

Male *Cftr* ^{tm1UNC}TgN(FABP*CFTR*) are bitransgenic mice harboring the FABP-h*CFTR* transgene [rat fatty acid binding protein 2, intestinal promoter directing expression of a human *CFTR* (ATP-binding cassette sub-family C, member 7) gene] and a targeted KO mutation of the *Cftr* gene.

(*Cftr* ^{tm1UNC}TgN(FABPCFTR) (named KO mice) and their wild-type congenic mice (named WT mice) (12–19 weeks of age) were maintained individually in single sterile cages in specific pathogen-free conditions in a ventilated isolator for two weeks before infection to consent acclimatization and to avoid microbiome perturbations. Mice were anesthetized by an i.p injection of 2.5% Avertin (2,2,2-Tribromoethanol, Sigma-Aldrich, Merck KGaA, Germany) and intratracheally infected with 3–4 × 10^5 of clinical strain RP73 *P. aeruginosa* [21], embedded in agar beads. A group of mice were challenged with sterile empty beads as a control. Mice were monitored daily for the healthy status and for the changes in body weight; the animals were sacrificed at seven days post-infection to track changes during long-term chronic infection as described in the experimental schedule (Figure 5). After the sacrifice, lungs, stool and the final straight portion of the rectum (mucosae + feces) were aseptically excised. Lungs were homogenized and split into aliquots for bacterial load count and microbiota analysis. An aliquot of 200 µL was immediately plated for bacterial evaluation, 500 µL was centrifuged and the supernatant was collected and frozen at −80 °C for additional analysis. In the remaining aliquot ATL buffer (15% of total volume, Qiagen, Germany) was added and stored at −80 °C for DNA extraction and metagenomic investigations. Stool and rectum were frozen at −80 °C to evaluate the composition of gut microbiome. A total of four experiments were performed. Statistical analyses were performed with GraphPad Prism (GraphPad Software, Inc., San Diego, CA, USA) using a two-way ANOVA with Bonferroni's multiple comparison test for body weight changes and Mann Whitney *t*-test for the other readouts. Outlier data, identified by Grubbs' test, were excluded from the analysis.

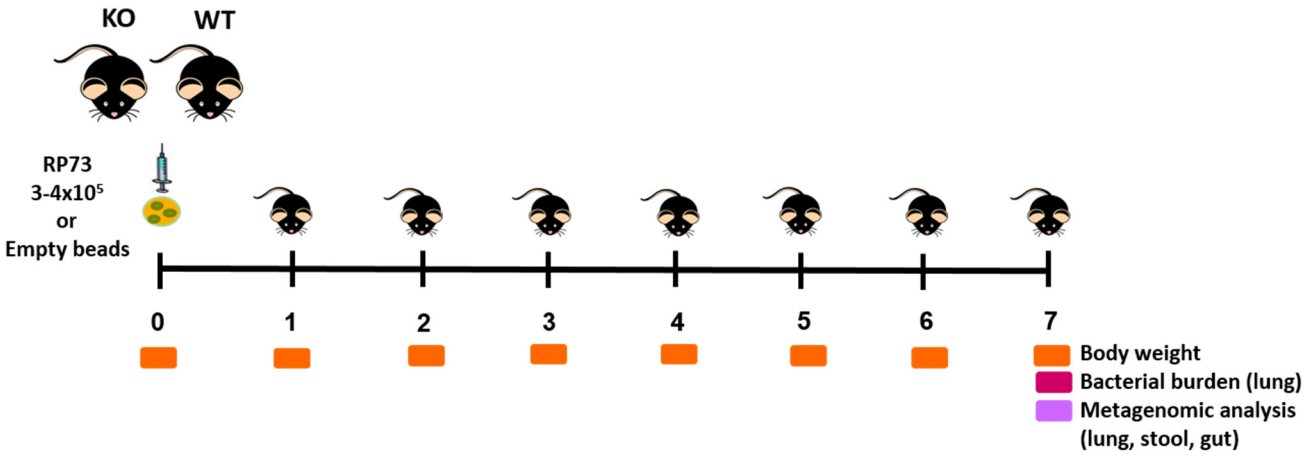


Figure 5. Experimental design. Mice were intratracheally infected with 3–4 × 10^5 of clinical strain *P. aeruginosa* RP73 embedded in agar beads or empty beads and were sacrificed after seven days. Lungs, stool, and the final straight portion of the rectum (mucosae + feces) were aseptically excised and processed for metagenomic analysis. Read-outs of the disease progression were body weight changes, health status and CFUs in the lung.

4.3. DNA Extraction and Sequencing

To set-up the best methodology of DNA extraction from mice tissues (lung homogenates, gut, and stool) stored at −80 °C, different methods were considered. After some preliminary tests, DNeasy PowerSoil Pro Kit (Qiagen, Germany) was selected for the stool and gut samples. Frozen stool and gut samples stored at −80 °C were added to PowerBead Pro tubes and treated as described in the manufacturer's instructions. DNA was recovered in 10mM Tris pH 7.4 and quantified using Qubit 2.0 (Thermo Fisher Scientific, MA, USA) fluorometer as per manufacturer's instructions. For lung homogenates, samples were subjected to DNA extraction by using a Qiagen DNeasy Blood and Tissue Kit, following the procedure reported by [34] with minor modification. Samples were subjected to bead beating for 1 min in DNA isolation bead tubes (Qiagen, Germantown, MD, USA) prior to kit use. The DNeasy Blood and Tissue protocol was modified to use 40 µL proteinase K instead of the recommended 20 µL, and samples were eluted with 50 µL of buffer AE instead of the suggested 200 µL. A total of 150 samples were collected (Table S2). The V3–V4 hypervariable region (approximately 469 bp) of the 16S rRNA gene, which has been used for taxonomic classification of the microbial community in human microbiome studies, was amplified. DNA was then sequenced according with the MiSeq platform (Illumina). Sequences are deposited under SRA PRJNA700486.

4.4. Bioinformatic and Statistical Analyses

The DADA2 pipeline (version 1.10.0) was used to reconstruct amplicon sequence variants (ASVs) from Illumina reads. Both ASV reconstruction and statistical analyses were performed in the R environment version 3.4.3 (http://www.R-project.org accessed on 1 September 2021). For a complete description of all the steps performed see the Supplementary Materials section. Briefly, primers used for V3-V4 amplification were detected and removed using cutadapt version 1.15. Low quality reads were discarded using the filterAndTrim function with an expected error threshold of two for both forward and reverse read pairs. Denoising was performed using the dada function after error rate modelling (learnErrors function). Denoised reads were then merged discarding those with any mismatches and/or an overlap length shorter than 20bp ('mergePairs' function). Chimeric sequences were removed using the removeBimeraDenovo function. Taxonomical classification was performed using DECIPHER package version 2.10.2 against the latest version of the pre-formatted Silva small-subunit reference database (SSU version 132. Available online: http://www2.decipher.codes/Downloads.html (accessed on 23 November 2020)). All sequences classified as chloroplasts, mitochondria, Archaea and Eukarya were

removed. Bacterial diversity in each sample was computed using the inverse Simpson index as implemented in the diversity function of the vegan package. Distances across samples were reported using non-metric multidimensional scaling (nMDS) as implemented in the metaMDS function of the vegan package, with 300 random starts and monotone regression. Differential abundance analysis was performed using the R package DESeq2 v1.30.0 [35].

5. Conclusions

Overall, our finding indicates that the *P. aeruginosa* infection influences murine health both in KO and WT mice significantly changing body weight both in CF and WT mice. The pulmonary chronic infection in CF mice may lead to altered intestinal mucosa not directly related to the CFTR genotype. Our results reinforce the hypothesis of a cross-talk between lung and gut microbiota in relation to *P. aeruginosa* chronic infection and CFTR mutation and emphasize the need of animal models to mechanistically decipher the influence of infection on whole host physiology. Results from such models, though oversimplified with respect to human patients, can give us a set of tools to unlock the potential of microbiome-based personalized medicine in major disease areas including CF. Moreover, such animal models will also assist in the development of therapeutic treatments which while targeting a specific microbiome (e.g., the gut microbiome) produce pleiotropic effect on other relevant organs (e.g., lung) as we may envisage from the presented results on a gut-to-lung axis in CF mice.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijms222212169/s1.

Author Contributions: Conceptualization, A.B. (Annamaria Bevivino), A.B. (Alessandra Bragonzi), and A.M.; methodology, A.B. (Annamaria Bevivino), G.B., A.B. (Alessandra Bragonzi), N.S., A.M.; software, G.B., L.C.; validation, G.B., A.M., A.B. (Annamaria Bevivino); formal analysis, G.B., A.R., I.D.F., F.A., L.C.; investigation, A.B. (Annamaria Bevivino), A.B. (Alessandra Bragonzi), A.M., G.B., L.C, F.A., A.R., I.D.F., N.S.; resources, A.B. (Annamaria Bevivino), A.B. (Alessandra Bragonzi), A.M.; data curation, A.B. (Alessandra Bragonzi), G.B., L.C, F.A., A.R.; writing—original draft preparation, A.B. (Annamaria Bevivino), A.B. (Alessandra Bragonzi), A.M., G.B., L.C.; writing—review and editing, A.B. (Annamaria Bevivino), A.B. (Alessandra Bragonzi), A.M., G.B., L.C, F.A., A.R., I.D.F., N.S.; visualization, L.C., G.B.; supervision, A.B. (Annamaria Bevivino); project administration, A.B. (Annamaria Bevivino); funding acquisition, A.B. (Annamaria Bevivino). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Italian Cystic Fibrosis Research Foundation, Grant FFC#19/2017 with the contribution of "Delegazione FFC Lago di Garda con i Gruppi di Sostegno FFC di Chivasso, dell'Isola Bergamasca, di Arezzo". G.B. was supported by a postdoctoral fellowship by the Italian Cystic Fibrosis Research Foundation.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, adhered to the Italian Ministry of Health guidelines for the use and care of experimental animals (IACUC 908), and was approved by the Institutional Review Board of IRCCS San Raffaele Scientific Institute.

Informed Consent Statement: Not applicable.

Data Availability Statement: Sequences reads are deposited at the National Center for Biotechnology Information (NCBI) database under the Bioproject PRJNA700486.

Acknowledgments: We dedicate this study to the memory of the late Professor Gianni Mastella, scientific director and co-founder of the Italian Cystic Fibrosis Research Foundation (FFC). On his death, we lost an excellent scientist, a loyal and generous friend, a wonderful father, and a marvelous speaker, a charming person of the highest sensitivity and integrity.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Gibson, R.L.; Burns, J.L.; Ramsey, B.W. Pathophysiology and Management of Pulmonary Infections in Cystic Fibrosis. *Am. J. Respir. Crit. Care Med.* **2003**, *168*, 918–951. [CrossRef]
2. Haack, A.; Aragão, G.G.; Novaes, M.R.C.G. Pathophysiology of Cystic Fibrosis and Drugs Used in Associated Digestive Tract Diseases. *World J. Gastroenterol.* **2013**, *19*, 8552–8561. [CrossRef]
3. Françoise, A.; Héry-Arnaud, G. The Microbiome in Cystic Fibrosis Pulmonary Disease. *Genes* **2020**, *11*, 536. [CrossRef]
4. Thavamani, A.; Salem, I.; Sferra, T.J.; Sankararaman, S. Impact of Altered Gut Microbiota and Its Metabolites in Cystic Fibrosis. *Metabolites* **2021**, *11*, 123. [CrossRef]
5. Zhang, D.; Li, S.; Wang, N.; Tan, H.-Y.; Zhang, Z.; Feng, Y. The Cross-Talk Between Gut Microbiota and Lungs in Common Lung Diseases. *Front. Microbiol.* **2020**, *11*, 301. [CrossRef]
6. Fouhy, F.; Ronan, N.J.; O'Sullivan, O.; McCarthy, Y.; Walsh, A.M.; Murphy, D.M.; Daly, M.; Flanagan, E.T.; Fleming, C.; McCarthy, M.; et al. A Pilot Study Demonstrating the Altered Gut Microbiota Functionality in Stable Adults with Cystic Fibrosis. *Sci. Rep.* **2017**, *7*, 6685. [CrossRef] [PubMed]
7. Ubags, N.D.J.; Marsland, B.J. Mechanistic Insight into the Function of the Microbiome in Lung Diseases. *Eur. Respir. J.* **2017**, *50*, 1602467. [CrossRef]
8. Sokolowska, M.; Frei, R.; Lunjani, N.; Akdis, C.A.; O'Mahony, L. Microbiome and Asthma. *Asthma Res. Pract.* **2018**, *4*, 1. [CrossRef] [PubMed]
9. Aarnoutse, R.; Ziemons, J.; Penders, J.; Rensen, S.S.; de Vos-Geelen, J.; Smidt, M.L. The Clinical Link between Human Intestinal Microbiota and Systemic Cancer Therapy. *Int. J. Mol. Sci.* **2019**, *20*, 4145. [CrossRef] [PubMed]
10. Anand, S.; Mande, S.S. Diet, Microbiota and Gut-Lung Connection. *Front. Microbiol.* **2018**, *9*, 2147. [CrossRef]
11. Dickson, R.P.; Erb-Downward, J.R.; Falkowski, N.R.; Hunter, E.M.; Ashley, S.L.; Huffnagle, G.B. The Lung Microbiota of Healthy Mice Are Highly Variable, Cluster by Environment, and Reflect Variation in Baseline Lung Innate Immunity. *Am. J. Respir. Crit. Care Med.* **2018**, *198*, 497–508. [CrossRef] [PubMed]
12. Zhou, A.; Lei, Y.; Tang, L.; Hu, S.; Yang, M.; Wu, L.; Yang, S.; Tang, B. Gut Microbiota: The Emerging Link to Lung Homeostasis and Disease. *J. Bacteriol.* **2020**, *203*, e00454-20. [CrossRef] [PubMed]
13. Semaniakou, A.; Croll, R.P.; Chappe, V. Animal Models in the Pathophysiology of Cystic Fibrosis. *Front. Pharmacol.* **2019**, *9*, 1475. [CrossRef]
14. Bragonzi, A. Murine Models of Acute and Chronic Lung Infection with Cystic Fibrosis Pathogens. *Int. J. Med. Microbiol.* **2010**, *300*, 584–593. [CrossRef]
15. Ellsworth, R.E.; Jamison, D.C.; Touchman, J.W.; Chissoe, S.L.; Braden Maduro, V.V.; Bouffard, G.G.; Dietrich, N.L.; Beckstrom-Sternberg, S.M.; Iyer, L.M.; Weintraub, L.A.; et al. Comparative Genomic Sequence Analysis of the Human and Mouse Cystic Fibrosis Transmembrane Conductance Regulator Genes. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 1172–1177. [CrossRef]
16. Lansdell, K.A.; Delaney, S.J.; Lunn, D.P.; Thomson, S.A.; Sheppard, D.N.; Wainwright, B.J. Comparison of the Gating Behaviour of Human and Murine Cystic Fibrosis Transmembrane Conductance Regulator Cl- Channels Expressed in Mammalian Cells. *J. Physiol.* **1998**, *508*, 379–392. [CrossRef]
17. Zhou, L.; Dey, C.R.; Wert, S.E.; DuVall, M.D.; Frizzell, R.A.; Whitsett, J.A. Correction of Lethal Intestinal Defect in a Mouse Model of Cystic Fibrosis by Human CFTR. *Science* **1994**, *266*, 1705–1708. [CrossRef] [PubMed]
18. Norkina, O.; Burnett, T.G.; De Lisle, R.C. Bacterial Overgrowth in the Cystic Fibrosis Transmembrane Conductance Regulator Null Mouse Small Intestine. *Infect. Immun.* **2004**, *72*, 6040–6049. [CrossRef]
19. Bazett, M.; Bergeron, M.-E.; Haston, C.K. Streptomycin Treatment Alters the Intestinal Microbiome, Pulmonary T Cell Profile and Airway Hyperresponsiveness in a Cystic Fibrosis Mouse Model. *Sci. Rep.* **2016**, *6*, 19189. [CrossRef] [PubMed]
20. Bazett, M.; Honeyman, L.; Stefanov, A.N.; Pope, C.E.; Hoffman, L.R.; Haston, C.K. Cystic Fibrosis Mouse Model-Dependent Intestinal Structure and Gut Microbiome. *Mamm. Genome* **2015**, *26*, 222–234. [CrossRef]
21. Facchini, M.; De Fino, I.; Riva, C.; Bragonzi, A. Long Term Chronic Pseudomonas Aeruginosa Airway Infection in Mice. *J. Vis. Exp.* **2014**, *85*, e51019. [CrossRef]
22. Cigana, C.; Lorè, N.I.; Riva, C.; De Fino, I.; Spagnuolo, L.; Sipione, B.; Rossi, G.; Nonis, A.; Cabrini, G.; Bragonzi, A. Tracking the Immunopathological Response to Pseudomonas Aeruginosa during Respiratory Infections. *Sci. Rep.* **2016**, *6*, 21465. [CrossRef]
23. Marsland, B.J.; Gollwitzer, E.S. Host–Microorganism Interactions in Lung Diseases. *Nat. Rev. Immunol.* **2014**, *14*, 827–835. [CrossRef]
24. Enaud, R.; Prevel, R.; Ciarlo, E.; Beaufils, F.; Wieërs, G.; Guery, B.; Delhaes, L. The Gut-Lung Axis in Health and Respiratory Diseases: A Place for Inter-Organ and Inter-Kingdom Crosstalks. *Front. Cell. Infect. Microbiol.* **2020**, *10*, 9. [CrossRef] [PubMed]
25. Huazano-García, A.; Shin, H.; López, M.G. Modulation of Gut Microbiota of Overweight Mice by Agavins and Their Association with Body Weight Loss. *Nutrients* **2017**, *9*, 821. [CrossRef] [PubMed]
26. Angelakis, E.; Armougom, F.; Million, M.; Raoult, D. The Relationship between Gut Microbiota and Weight Gain in Humans. *Future Microbiol.* **2012**, *7*, 91–109. [CrossRef] [PubMed]
27. Vernocchi, P.; Del Chierico, F.; Russo, A.; Majo, F.; Rossitto, M.; Valerio, M.; Casadei, L.; La Storia, A.; De Filippis, F.; Rizzo, C.; et al. Gut Microbiota Signatures in Cystic Fibrosis: Loss of Host CFTR Function Drives the Microbiota Enterophenotype. *PLoS ONE* **2018**, *13*, e0208171. [CrossRef]

28. Meeker, S.M.; Mears, K.S.; Sangwan, N.; Brittnacher, M.J.; Weiss, E.J.; Treuting, P.M.; Tolley, N.; Pope, C.E.; Hager, K.R.; Vo, A.T.; et al. CFTR Dysregulation Drives Active Selection of the Gut Microbiome. *PLoS Pathog.* **2020**, *16*, e1008251. [CrossRef]
29. Groves, H.T.; Cuthbertson, L.; James, P.; Moffatt, M.F.; Cox, M.J.; Tregoning, J.S. Respiratory Disease Following Viral Lung Infection Alters the Murine Gut Microbiota. *Front. Immunol.* **2018**, *9*, 182. [CrossRef] [PubMed]
30. O'Toole, G.A.; Crabbé, A.; Kümmerli, R.; LiPuma, J.J.; Bomberger, J.M.; Davies, J.C.; Limoli, D.; Phelan, V.V.; Bliska, J.B.; DePas, W.H.; et al. Model Systems to Study the Chronic, Polymicrobial Infections in Cystic Fibrosis: Current Approaches and Exploring Future Directions. *MBio* **2021**, *12*, e0176321. [CrossRef]
31. Bevivino, A.; Bacci, G.; Drevinek, P.; Nelson, M.T.; Hoffman, L.; Mengoni, A. Deciphering the Ecology of Cystic Fibrosis Bacterial Communities: Towards Systems-Level Integration. *Trends Mol. Med.* **2019**, *25*, 1110–1122. [CrossRef]
32. Gabrielaite, M.; Johansen, H.K.; Molin, S.; Nielsen, F.C.; Marvig, R.L. Gene Loss and Acquisition in Lineages of Pseudomonas Aeruginosa Evolving in Cystic Fibrosis Patient Airways. *MBio* **2020**, *11*, e02359-20. [CrossRef]
33. Freschi, L.; Vincent, A.T.; Jeukens, J.; Emond-Rheault, J.-G.; Kukavica-Ibrulj, I.; Dupont, M.-J.; Charette, S.J.; Boyle, B.; Levesque, R.C. The Pseudomonas Aeruginosa Pan-Genome Provides New Insights on Its Population Structure, Horizontal Gene Transfer, and Pathogenicity. *Genome Biol. Evol.* **2019**, *11*, 109–120. [CrossRef] [PubMed]
34. Mason, K.L.; Downward, J.R.E.; Mason, K.D.; Falkowski, N.R.; Eaton, K.A.; Kao, J.Y.; Young, V.B.; Huffnaglea, G.B. Candida Albicans and Bacterial Microbiota Interactions in the Cecum during Recolonization Following Broad-Spectrum Antibiotic Therapy. *Infect. Immun.* **2012**, *80*, 3371–3380. [CrossRef] [PubMed]
35. Love, M.I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef] [PubMed]

Synergistic effect of rhizobia and plant-growth promoting nonsymbiotic bacteria on alfalfa growth

Bianco C., Cangioli L., Fagorzi C., Amato G., Mengoni A., Defez R.

Abstract

One of the main strengths of the sustainable agriculture is the biological nitrogen fixation to reduce the input of chemical fertilizers in agriculture systems. The use of biofertilizers meets the goal to reduce the negative impacts that chemical fertilizers have on soil and water and an improvement of the symbiotic activity of leguminous plants in now compulsory. The aim of the MICRO4Legumes project is to increase the positive effects that nitrogen-fixing rhizobia bring to the fodder legumes cultivated in biological agricultural systems, exploiting the functional and genetic biodiversity of several rhizobia strains naturally present in the soil. Although the scientific community considers bacterial nitrogen fixation of enormous importance, it still remains to be clarified which rhizobial consortia are the most effective in different pedo-climatic environments. It is in fact known that different strains of rhizobia, even of the same species, have significantly different effects depending on the species, cultivar, soil conditions, and agronomic and environmental conditions. It is also known that the association of these rhizobia with non-rhizobial soil bacteria can improve the growth of fodder legumes. Therefore, it is essential to evaluate the different rhizobial consortia in specific environments to determine the most effective one in improving the yield and quality of the forage obtained. Moreover, it is necessary to identify the non-rhizobial partners to be used for co-inoculation. In particular, we selected rhizobial strains able to improve the yield and the resilience of alfalfa in harsh conditions. After the creation of new microbial consortia, the most effective have been in-vitro tested and non-rhizobial bacterial strains have been selected to be co-inoculated with rhizobial consortia. Subsequently, microbial consortia were directly applied in field and we measured the physiological response of plants through agronomic tests. Finally, we compared the soil and the rhizospheric microbiome before and after the inoculation.

Introduction

The concept of modern and sustainable agriculture implies the exploitation of the biological nitrogen fixation (BNF). The nitrogen input coming from the BNF allows us to reduce the usage of chemical nitrogen fertilizers which negatively impact water and soil health. As imminent challenges, sustainable agriculture should improve the utilization of symbiotic microorganisms in fodder crops, as the enactment of environmental restrictions[1]. The main purpose of this project aims to satisfy these needs of environmental sustainability increasing the beneficial effects of nitrogen-fixing rhizobia concerning the crop yield and quality of fodder legumes in biological agriculture. To achieve this goal, it is necessary to better decipher the interactions between plant and symbiotic microorganisms to develop smart pharming and precision agriculture strategies exploiting the genetic and functional diversity of native rhizobia present in the soil. Although scientific community considers the bacterial nitrogen fixation a relevant and promising approach, little is known about the most efficient rhizobial consortia in different pedo-climatic environments. Indeed, it is well known that different strains of rhizobia, even belonging to the same species, could have different impacts in relation to species, cultivar, soil features, and agronomic and environmental conditions. In addition, the association of these rhizobial strains with non-rhizobial soil bacteria could improve the fodder legumes growth. Thus, it becomes essential not only to evaluate and discover which are the most efficient rhizobial consortia for each specific environment, but also to detect non-rhizobial partners to use for the co-inoculation.

Alfalfa is the most widespread fodder crop cultivated in southern Italy and the main temperate forage legume crop in the world. Conventionally, alfalfa is cultivated without the need of inoculants, thanks to the

ability to form nitrogen-fixing symbiotic association with native rhizobia present in soil, allowing to growth the plant even in poor soil and marginal lands without the need for nitrogen fertilization input. However, naturally occurring rhizobia are often quite variable in their effect on plants, resulting in suboptimal symbiotic nitrogen-fixing performances[2,3]. The use of seed coated with selected elite rhizobia strains is a growing practice and now several seed producers and company of bioinoculant production are selling coated alfalfa seeds[4,5]. However, it is becoming clear that rhizobia alone can provide only a part of the plant growth promotion effect. The possibility to increase mineral nutrition by other component of the plant microbiota (as for instance mycorrhizal fungi) is stirring the attention toward novel microbial consortia, including several types of microbes and mimicking in part the native plant-associated microbiome, called synthetic communities (SynCom)[6,7].

Here we aimed at providing a proof-of-principle on the improved efficacy in plant growth promotion of a consortium of microbes composed by rhizobia and nonsymbiotic plant-associated bacteria (nonrhizobia), from previously established collection, having the ability to tolerance salt stress, produce phytohormones (as auxins), to fix atmospheric nitrogen in free-living conditions (for nonrhizobia) and endophytically colonize plant tissues for nonrhizobia. This research activity is included into the research project funded by the Italian Ministry of Agriculture *"Il microbioma vegetale simbionte come strumento per il miglioramento delle leguminose foraggere" (MICRO4Legumes)*

Methods

Strains and microbiological analyses. The selected rhizobial strains are ten (n=10) and they are listed in Table 1 with their specific characteristics.

Table 1. *Selected Sinorhizobium meliloti strains. When no indicated, strains were isolated from root nodules of alfalfa plants.*

Strain	Geographical origin	Genome sequence (genbank code)	Level of NaCl tolerance (concentration in TY medium)
AK58	Salinized area in northern Aral Sea	GCA_000473425.1	1 M
AK83	Salinized area in northern Aral Sea (from *Medicago falcata* root nodules)	GCA_000147795.3	1 M
BL225C	Italian agricultural soil	GCA_000147775.3	nd
BO21CC	Italian agricultural soil	GCA_000473405.1	1 M
H1	Italian agricultural soil (isolated from alfalfa leaves)	GCA_000287475.1	1 M
AO643DD	Italian agricultural soil	GCA_000287575.1	1 M
CO438LL	Italian agricultural soil	GCA_000287495.1	2 M
AK75	Salinized area in northern Aral Sea	GCA_000287515.2	2 M
CO431A	Italian agricultural soil	GCA_000287375.1	1 M
SM11	German agricultural soil	GCA_000218265.1	0.4 M

Auxin production. The physiological characterization of the strains involved the measurement of the levels of indole-3-acetic acid (IAA) auxin produced in liquid culture using a colorimetric method which involves the

use of the Salkowski reagent. Liquid cultures of each strain were treated for approximately 18 hours with tryptophan, the precursor of IAA. The used reagent contains iron and perchloric acid and allows the evaluation of the pink colour produced following the formation of a tris-(indole-3-acetate)-iron(III) complex. At least 5 independent cultures were established for each strain to be characterized.

Genome sequencing and annotation. From all selected strains, genomic DNA was extracted to perform the genome sequencing through the Pacific Biosciences Sequel platform (PacBio). The sequenced genomes were assembled and annotated with the aim to identify the potential genes involved in plant-growth promoting (PGP) traits and develop barcoding systems allowing to trace back the previously applied bioinoculants in the soil and in the plant.

In vitro test of the consortia. Cross inhibition tests between the selected strains were performed to verify the possible co-existence of all microorganisms present in the synthetic community. The tests were performed using the technique of aga diffusion, as previously described[8]. Not any cross inhibition was detected (data not shown). These results allow us to identify the combination of endophytes and rhizobia to apply as bioinoculants.

Field study. Previous investigations aimed to select the alfalfa cultivar to inoculate with synthetic consortia in field trials. Principally, three cultivars (Gea - Italia, Hunter River – Australia, NB Banat ZMSII - Serbia) with different geographical origin and drought tolerance were selected. Each cultivar was treated with different inocula: with the selected rhizobial strain of *Sinorhizobium meliloti* BL225C, with each single strain of selected endophytes and with the mix composed by the endophytes. Alfalfa plants were inoculated with the only rhizobial strain were used as negative control. The nitrogen fixing ability after the inoculation on plants was evaluated through an in-vitro test of nitrogenase enzyme activity using the acetylene as alternative to atmospheric nitrogen through gas chromatography.

Microbial consortia efficiency. After preliminary in-vitro experiments, the Mix1 and the Mix2 inocula were directly tested in field in the following three biological farms:

- **Azienda Leone Alberto**
  Via Laghetto/C,da Sellino, 84047 Capaccio Paestum (SA), Campania,
- **Azienda Tedesco Anna Maria**
  Via Cesare Battisti n, 35, 82024 Colle Sannita (BN)
- **Azienda Paolucci Lina**
  C/da Marcandindoli SNC, 82024 Colle Sannita (BN), Campania

Sowing trials were conducted on parcels of 250 m^2 previously examined and selected. The experimental scheme was composed by 4 randomized blocks with 4 repetitions and single parcels of 10 m^2 (Figure 2),

Analysis of rhizospheric microbial communities. For each of the six blocks, a total of 9 samples was collected (3 inocula x 3 irrigation levels) for a total of 54 samples. After digging up the entire plant (10 cm radius around the plant and down to 20 cm soil depth) and shaking off the plant roots, the soil remaining attached to the roots was detached with a sterile scalpel and considered to be rhizospheric soil. Soil samples were collected in triplicate at a depth of 15 cm, and then mixed.

Environmental DNA (eDNA) was extracted from 250 mg of soil using the DNeasy PowerSoil Pro kit (Qiagen). From the extracted DNA, the bacterial V4 region of 16S rRNA genes was amplified with primers 515F (5'-GTGCCAGCMGCCGCGGTAA-3') and 806R (5'-GACTACHVGGGTATCTAATCC-3') following the protocol describer in our previous works[9–11]. PCR products were sequenced in a single run using Illumina MiSeq technology

with pair-end sequencing strategy and a MiSeq Reagent Kit v3 (Illumina, USA). Library preparation and demultiplexing were performed following Illumina's standard pipeline as previously reported. Illumina reads were processed following the validated pipeline reported in previous works[9–11]. The analysis of microbial communities was performed through the pipeline described in Cangioli et al., 2022[9–11].

Results

Rhizobial strains selection. First of all, a screening of available strains from already established collections was carried out to select the best candidates for the creation of the microbial consortium in relation to their already studied beneficial properties. Among the ten strains of alfalfa nitrogen-fixing rhizobium *S. meliloti* selected, priority was given to strains an isolated from different geographical areas, including the desertified and salinized area of Aral Sea, with the purpose to select strains able to tolerate arid soil conditions.

Concerning the non-rhizobial strains able to promote alfalfa growth, a collection of endophytic bacterial strains previously isolated from different plant species (as rice) and tested for their plant colonization and growth abilities. These bacterial strains belong to β- and ϒ-Proteobacteria species known for their plant association trait.

Table 2. *Selected endophytic strains[9].*

Taxonomic affiliation	Code
Citrobacter sp.	BDA59-3
Klebsiella sp.	BDA62-2
Enterobacter sp.	BDA62-3
Ralstonia sp.	BDA134-6
Microbacterium sp.	BDA137-13
Enterobacter sp.	RCA25
Herbaspirillum sp.	RCA24

Strains and consortia physiological properties tests. The physiological characterization was performed for seventeen (n=17) bacterial strains.

The results obtained and reported in Table 3. Approximately 80% of the rhizobium strains produced low and very similar levels of IAA, while only two strains (BM499 and BM508) produced greater quantities of auxin. On the other hand, endophytic strains returned an opposite result (highlighted in grey in Table 3): approximately 70% of them produced good IAA levels, while for the remaining 30% levels very similar to those recorded for the majority of rhizobia were measured. The analysis of the results obtained will allow us to identify a subset of strains to produce a minimal synthetic community to be tested in alfalfa. In particular, strains BL225C, BDA59-3, BDA62-2, BDA62-3, BDA134-6, BDA137-13, RCA25, RCA24 were selected. The choice of *S. meliloti* BL225C strain relied on previous results also, which indicate this strain as highly effective in symbiotic nitrogen fixation and in competition with other strains[12,13].

Table 3. *Evaluation of the indole-3-acetic acid (IAA) auxin production.*

Strain	IAA production (nmol IAA/mg cells)
Sinorhizobium meliloti BM492	4.7 ± 0.6
Sinorhizobium meliloti BM493	4.6 ± 0.4
Sinorhizobium meliloti BM494	6.5 ± 0.9
Sinorhizobium meliloti BM495	4.7 ± 0.4
Sinorhizobium meliloti BM498	3.7 ± 0.2
Sinorhizobium meliloti BM499	13 ± 3
Sinorhizobium meliloti BM501	6.5 ± 0.5
Sinorhizobium meliloti BM504	3.3 ± 0.2
Sinorhizobium meliloti BM506	3.5 ± 0.3
Sinorhizobium meliloti BM805	12 ± 1
Citrobacter sp. BDA59-3	61 ± 10
Klebsiella sp.BDA62-2	56 ± 12
Enterobacter sp. BDA62-3	7.7 ± 0.5
Ralstonia sp.BDA134-6	31 ± 2
Microbacterium sp. BDA137-13	17 ± 2
Enterobacter sp. RCA25	3.1 ± 0.3

Herbaspirillum sp. RCA24	36 ± 11

Genomic analysis of bacterial strains. The majority of the genomes of selected strains used in this work were already sequences, assembled and annotated in Bianco et al., 2021[9]. Genomes of RCA24 and RCA 25 strains were later sequenced through PacificBioscience platform (PacBio) at the Department of Biology at the University of Florence (Italy). Genomes features are reported in Table 4.

Table 4. *Genomes analysis of RCA24 and RCA25 strains.*

	RCA24	**RCA25**
N. of replicons	2	4
Genome size	4731967	5080969
GC content %	55.84	53.66
Best taxonomic match	*Enterobacter chengduensis/ Enterobacter asburiae*	*Kosakonia sacchari*

Creation and testing of synthetic microbial consortia. Two consortia were prepared: Mix1 which includes one rhizobium strain (BL225C9 and two IAA-producing strains (BDA62-2 and BDA137-13); Mix2, which is composed by the same rhizobium strain and include all other nonsymbiotic endophytes. Both consortia contained 10^6 cells/ml per each strain (for the rhizobium BL225C 10^4 cells/ml) (Table 5).

Table 5. *List of the strains (rhizobia and endophytes) contained in the inoculation mixes.*

Mix1 Inoculum	**Mix2 Inoculum**
BL225C	BL225C
BDA62-2	BDA59-3
BDA137-13	BDA62-2
	BDA62-3
	BDA134-6
	BDA137-13
	RCA24
	RCA25

Impact of bioinoculant consortia on soil microbiome. As reported in methods, a total of 54 samples was collected (9 samples for each of the six blocks). In particular, for the 16S rRNA gene sequencing, a total of 27 samples was analysed (3 inocula x 3 irrigation levels x 3 replicates). In Table 6, a list of all analysed samples is reported.

Table 6. *List and description of analysed samples.*

sample	***tesi***	***irrigazione***
*ctr_100pc_*b1	*Ctr*	*100_pc*
*ctr_50pc_*b1	*Ctr*	*50_pc*
*ctr_75pc_*b1	*Ctr*	*75_pc*
*mix1_100pc_*b1	*Mix_1*	*100_pc*
*mix1_50pc_*b1	*Mix_1*	*50_pc*
*mix1_75pc_*b1	*Mix_1*	*75_pc*
*mix2_100pc_*b1	*Mix_2*	*100_pc*
*mix2_50pc_*b1	*Mix_2*	*50_pc*
*mix2_75pc_*b1	*Mix_2*	*75_pc*
*ctr_100pc_*b2	*Ctr*	*100_pc*
*ctr_50pc_*b2	*Ctr*	*50_pc*
*ctr_75pc_*b2	*Ctr*	*75_pc*
*mix1_50pc_*b2	*Mix_1*	*50_pc*
*mix1_75pc_*b2	*Mix_1*	*75_pc*
*mix1_100pc_*b3	*Mix_1*	*100_pc*
*mix1_50pc_*b3	*Mix_1*	*50_pc*
*mix2_50pc_*b3	*Mix_2*	*50_pc*
*mix2_75pc_*b3	*Mix_2*	*75_pc*
*ctr_100pc_*b4	*Ctr*	*100_pc*
*mix1_100pc_*b3_1	*Mix_1*	*100_pc*
*mix2_100pc_*b1_1	*Mix_2*	*100_pc*
*mix2_50pc_*b1_1	*Mix_2*	*50_pc*

*mix2_75pc*_b1_1	*Mix_2*	*75_pc*
*ctr_50pc*_b2_1	*Ctr*	*50_pc*
*ctr_75pc*_b2_2	*Ctr*	*75_pc*
*mix1_75pc*_b2_1	*Mix_1*	*75_pc*
*mix1_100pc*_b1_2	*Mix_2*	*100_pc*

After the trimming of adaptors sequence, raw sequences obtained by the sequencing were processed following the standard DADA2 pipeline[14]. The taxonomic assignment to the Amplicon Sequence Variants (ASVs) was performed using the Silva database (https://www.arb-silva.de/).

A satisfactory number of reads was obtained from results of 16S rRNA gene libraries for most of samples, which after analysis of amplicon sequence variants allowed to reach nearly a saturation in a rarefaction analysis (Figure 1) and Goods' coverage values above 0.99. Samples with a number of reads lower that 10'000 were discarded.

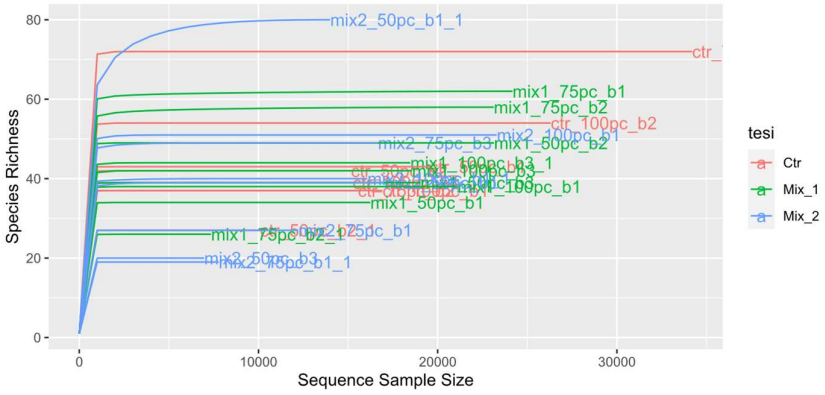


Figure 1. *Rarefaction curves of soil samples.*

Concerning the taxonomic distribution of the amplicon sequence variations (ASVs), those associated with Actinobacteriota, Proteobacteriota, Firmicutes, Verrucomicrobiota, Chloroflexi and Acidobacteriota were found to be the most abundant (alpha <0.05) (Figure 2).

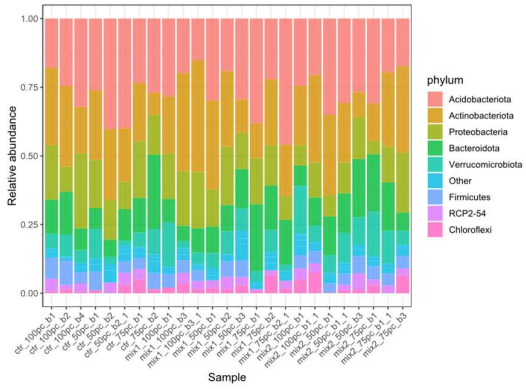


Figure 2. *Taxonomic composition of alfalfa microbiota.*

Alfa diversity of plants microbiota was performed for both two groups of treatments and irrigation levels by measuring the Shannon, Simpson and Richness indices. Different levels of irrigation and the application of the Mix1 and Mix2 inoculum have no influence on alpha diversity (p.value>0.05).

Table 7. *Wilcoxon test performed on alpha diversity analysis. Values refer to the mean and the standard deviation (SD).*

	Observed		Shannon		Simpson	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
ctr_100pc	43,33	9,29	3.49	0.19	0.96	0.01
ctr_50pc	37,33	8,96	3.33	0.26	0.96	0.01
ctr_75pc	54,50	24,75	3.64	0.50	0.96	0.02
mix1_100pc	41,33	3,06	3.36	0.10	0.96	0.01
mix1_50pc	40,67	7,64	3.41	0.22	0.96	0.01
mix1_75pc	48,67	19,73	3.46	0.61	0.95	0.04
mix2_100pc	45,50	7,78	3.53	0.22	0.96	0.01
mix2_50pc	46,33	30,66	3.19	0.40	0.94	0.02
mix2_75pc	31,67	15,53	3.10	0.42	0.94	0.02

Clustering of samples due to their taxonomic composition was performed through a nonMetric Multidimensional Scaling (NMDS) (Figure 3). A possible patterning is related to the Mix1 inoculum but PerMANOVA analysis showed no significant differences between the Control, the Mix1 and the Mix2 inoculum (p.value>0.05) (Table 8).

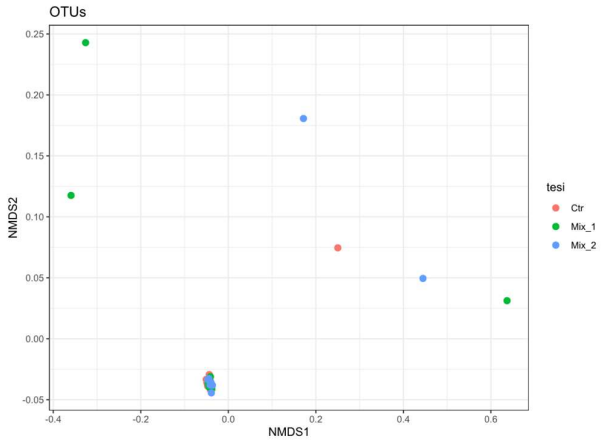


Figure 3. nonMetric Multidimensional Scaling of alfalfa samples, divided by treatment.

Table 8. *PerMANOVA test on beta-diversity performed among treatments and irrigation levels. Values refer to p.values.*

Permanova						
Treatment	**Df**	**SumsofSqs**	**F.Model**	**R2**	**p.value**	**p.adjusted**
Ctr vs Mix 1	1	0.480894	0.963845	0.060377	0.91	1
Ctr vs Mix 2	1	0.500359	1,0046398	0.066955	0.437	1
Mix 1 vs Mix 2	1	0.486299	0.976035	0.061094	0.78	1
Irrigation	**Df**	**SumsofSqs**	**F.Model**	**R2**	**p.value**	**p.adjusted**
100pc vs 50pc	1	0.502419	1,012117	0.063209	0.359	1
100pc vs 75pc	1	0.504576	1,014765	0.067584	0.323	0.969
50pc vs 75pc	1	0.515239	1,037550	0.064695	0.092	0.276

Conclusions

The data obtained so far provide proof-of-principle evidence that a microbial consortium composed by rhizobia and other plant growth promoting bacteria (PGPB) is more effective than the sole rhizobium application. Preliminary field trial then indicated a negligible effect of the inoculant application on the native rhizospheric soil microbiome, suggesting that inoculant application does not alter community structure and is not a perturbing (viz. alien) element for the soil microbiota. Future analyses are required to assess the effect of the consortia over in field alfalfa growth and possibly tolerance to drought conditions.

Bibliography

1. Abd-Alla, M. H., Al-Amri, S. M. & El-Enany, A.-W. E. Enhancing Rhizobium–Legume Symbiosis and Reducing Nitrogen Fertilizer Use Are Potential Options for Mitigating Climate Change. *Agriculture* **13**, 2092 (2023).
2. Kawaka, F. *et al.* Symbiotic Efficiency of Native Rhizobia Nodulating Common Bean (*Phaseolus vulgaris L.*) in Soils of Western Kenya. *Int. Sch. Res. Not.* **2014**, 1–8 (2014).
3. Bogino, P., Nievas, F., Banchio, E. & Giordano, W. Increased competitiveness and efficiency of biological nitrogen fixation in peanut via in-furrow inoculation of rhizobia. *Eur. J. Soil Biol.* **47**, 188–193 (2011).
4. Trabelsi, D. & Mhamdi, R. Microbial Inoculants and Their Impact on Soil Microbial Communities: A Review. *BioMed Res. Int.* **2013**, 863240 (2013).
5. Nelson, M. S. & Sadowsky, M. J. Secretion systems and signal exchange between nitrogen-fixing rhizobia and legumes. *Front. Plant Sci.* **6**, (2015).
6. Vaccaro, F., Cangioli, L., Mengoni, A. & Fagorzi, C. Synthetic plant microbiota challenges in nonmodel species. *Trends Microbiol.* **30**, 922–924 (2022).
7. Moon, T. S. SynMADE: synthetic microbiota across diverse ecosystems. *Trends Biotechnol.* **40**, 1405–1414 (2022).
8. Bonev, B., Hooper, J. & Parisot, J. Principles of assessing bacterial susceptibility to antibiotics using the agar diffusion method. *J. Antimicrob. Chemother.* **61**, 1295–1301 (2008).
9. Bianco, C. *et al.* Endophytes from African Rice (Oryza glaberrima L.) Efficiently Colonize Asian Rice (Oryza sativa L.) Stimulating the Activity of Its Antioxidant Enzymes and Increasing the Content of Nitrogen, Carbon, and Chlorophyll. *Microorganisms* **9**, 1714 (2021).
10. Cangioli, L. *et al.* Effect of Site and Phenological Status on the Potato Bacterial Rhizomicrobiota. *Microorganisms* **10**, 1743 (2022).

11. Cangioli, L. *et al.* Differential Response of Wheat Rhizosphere Bacterial Community to Plant Variety and Fertilization. *Int. J. Mol. Sci.* **23**, 3616 (2022).
12. Bellabarba, A. *et al.* Competitiveness for Nodule Colonization in Sinorhizobium meliloti: Combined *In Vitro* -Tagged Strain Competition and Genome-Wide Association Analysis. *mSystems* **6**, e00550-21 (2021).
13. Giuntini, E. *et al.* Large-scale genetic variation of the symbiosis-required megaplasmid pSymA revealed by comparative genomic analysis of Sinorhizobium meliloti natural strains. *BMC Genomics* **6**, 158 (2005).
14. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

Study of the regulatory cascade involved in the megaplasmids conjugation in the model plant symbiotic bacterium Sinorhizobium meliloti by the creation of a GFP-based reporter system

Lisa Cangioli[1], M. Fini[1], G. Cambi[1], A. Mengoni[1], C. Fagorzi[1]

[1] Department of Biology, University of Florence, via Madonna del Piano 6, 50019 Sesto Fiorentino, Italy

Abstract

Sinorhizobium meliloti is a model plant symbiotic bacterium. Its genome is composed by a chromosome, and two megaplasmids, pSymA and pSymB, which carry genes related to the processes of symbiotic nitrogen fixation and growth in the rhizosphere, respectively. The conjugal transfer of the symbiotic megaplasmids pSymA in *S. meliloti* is highly regulated and repressed through a system named rct, composed by 7 key genes. However, to date the environmental signals which trigger the rct regulatory cascade are still unknown. Here, we aim to decipher the signals of rct activation. Promoters of rct genes were amplified and cloned upstream to green fluorescent protein (GFP) gene in the pOT2 plasmid. Derivative *S. meliloti* strains harbouring the recombinant pOT2 plasmids were tested under various inducing conditions, including plant compounds as root exudate or plant macerate and acceptor strains lacking the megaplasmid pSymA. The number of cells in the population expressing GFP was counted through spectrophotometric reading in preliminary induction tests, and, consequently, through cytofluorimetric readings. Results showed that the fraction of cells expressing GFP was around 10^{-4} over the total cells counted. One promoter was found to be constitutively expressed, but no differences in the fraction of positive cells was found under the tested conditions. Finally, the pTE3rctB plasmid was transferred into selected recombinant *S. meliloti* strains, which were then tested. However, the developed panel of recombinant strains allows to plan further investigation under in vitro and in vivo (e.g. soil, plant symbiosis) conditions to identify the ecological determinants of symbiotic megaplasmid transfer in *S. meliloti*.

Methods

Except for the strains and plasmids constructed in this study, all other strains and plasmids have been previously described and are listed in Table 1.

Table 1. *Strains and Plasmid Used in This Study**

Strain or Plasmid	**Description**	**Reference**
Sinorhizobium meliloti		
1021	Wild type; Sm^R	Meade et al., 1982[1]
Escherichia coli		
DH5α	Designed by. Strain with maximised transformation efficiency.	Hanahan et al., 1991[2]
MT616	Helper strain carrying pRK2013; Cm^R	Finan et al., 1986[3]
Plasmids		
pOT2	Broad host range vector containing GFPuv promoterless; Gm^R	Karunakaran et al., 2005[4]
pTE3rctB	Broad host range vector overexpressing the *Rhizobium etli* rctB gene; Tc^R	Nogales et al., 2013[5]

*Code of strains and plasmid is reported. A succinct description of the main phenotypic features is shown; Sm^R, streptomycin resistance, Tc^R, tetracyclin resistance, Gm^R, gentamycin resistance.

Microbiological Methods. In Table 1, all strains and plasmids used in this study are described. *Sinorhizobium meliloti* strains were cultured on solid or liquid TY medium with 0.2g/l $CaCO_3$, while Escherichia coli strains were grown in Luria-Bertani (LB) medium, supplemented with antibiotics when necessary. All growth media (LB, TY, M9) and antibiotic concentrations were as previously described[6].

Primers selection. Primers were designed to amplify the promoters of the key genes that are involved in the regulatory cascade (Figure 1) which leads to the mobilization of pSymA megaplasmid, using the web software Primer3 (https://primer3.ut.ee/). For each pair of primers, specific tails containing restriction sites were added (Table 2).

Table 2. *List of primers.*

Code	Primer	Sequence	Tm (°C)
1	*p_traA2_SalI_F*	5' ttttgtcgacCGTTGATGGTATCGCTGTAG 3'	57
	p_traA2_PstI_R	5' ttttctgcagTGATGTCTCCGGACTTGTC 3'	57
3	*p_traA1_SalI_F*	5' ttttgtcgacGGCCTCGTTCTTGCGATCA 3'	64.7
	p_traA1_PstI_R	5' ttttctgcagTGATGTCTCCAGACTTGTCCAG 3'	60.3
5	*p_rctR_SalI_F*	5' ttttgtcgacATTCCGATGTTCCGGTGATT 3'	61.5
	p_rctR_PstI_R	5' ttttctgcagTTCCCTCTCTCACTCGGTTG 3'	58.9
6	*p_Sma0956-rctC_SalI_F*	5' ttttgtcgacGCTGAGGTCGAGAAACTGAAG 3'	59.2
	p_Sma0956-rctC_PstI_R	5' ttttctgcagAGTTCTGCTTCCCTGAGTGAG 3'	58.7
7	*p_rctB_SalI_F*	5' ttttgtcgacATAGTTGCGGCTTCGCTC 3'	63.8
	p_rctB_PstI_R	5' ttttctgcagGACCGCTCCTTATTGCTTCAC 3'	61.1
8	*p_virB/rctA_SalI_F*	5’ ttttgtcgacGGTCATGAACACGTCTCTCG 3’	58
	p_virB/rctA_BamHI_R	5’ ggatccGGCTCAGTACCACACCCTTT 3’	58
9	*p_virB/rctA_KpnI_F*	5' ttttggtaccGGCTCAGTACCACACCC 3'	54.5
	p_virB/rctA_SalI_R	5' ttttgtcgacCTTTTATACTGTCAGTTATATGAT 3'	47.3
	gfp_R	5' CTAATTCAACAAGAATTGGGACAAC 3'	57.1

BamHI	SalI	PstI	KpnI
ggatcc	ttttgtcgac	ttttctgcag	ttttggtacc

Recombinant strains construction. Genomic DNA (gDNA) extraction from *Sinorhizobium meliloti* 1021 was performed using the DNeasy PowerLyzer PowerSoil Kit (QIAGEN), following the manufacturer’s instructions. pOT2 plasmid was extracted from cultures of *Escherichia coli* DH5α using the QIAprep Spin Miniprep Kit (Qiagen), following the manufacturer’s instructions. The quality of gDNA and pOT2 plasmid samples and DNA quantification was analysed by 1% agarose gel electrophoresis. Selected promoters (listed in Table 2) were amplified using the DNA polymerase Phusion Flash II DNA Polymerase through PCR amplification (10 s 98°C, 30 x (1 s 98°C, 5 s 55°C, 10 s 72°C), 1 min 72°C). Purified pOT2 plasmid was subjected to two consequential restriction digestions with PstI and SalI/BamHI/KpnI. Restriction digestions were incubated overnight at 37 °C. After the first digestion, the enzyme was not inactivated by heat, as it was filtered with the kit QIAquick PCR Purification, whereas, after the second overnight incubation, the reaction was stopped heating at 60°C for 15 minutes. Subsequently, digested pOT2 plasmids were dephosphorylates using the

FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific). To purify DNA fragments from PCR and other enzymatic reactions, the QIAquick PCR Purification Kit or the QIAquick PCR & Gel Cleanup Kit was used. The DNA fragments containing promoters from 1021 strain were cloned into the promoter-less vector pOT2 containing GFP-uv (green fluorescent protein) as reporter gene[4] through an overnight ligation reaction at 14°C using T4 DNA Ligase (Thermo Scientific). Recombinant pOT2 plasmid was transformed into Ca^{2+}-competent cells of *Escherichia coli* DH5α, previously prepared as described in protocol[7]. Transformant cells were selected for resistances to gentamycin and verified by PCR amplification. A first triparental mating between the wild type strain 1021, the helper strain *E. coli* MT616 (carrying pRK600), and *E. coli* with the recombinant pOT2 vector[8]. Recombinant colonies were selected with the antibiotic resistances to gentamycin and streptomycin. A second triparental mating between the recombinant *S. meliloti* 1021 strains carrying the recombinant pOT2 plasmids, the helper strain *E. coli* MT616 and *E. coli* DH5α with the pTE3rctB vector (replicative plasmid overexpressing the *R. etli rctB* gene and carrying a tetracycline resistance marker)[5]. Selection of recombinant colonies of *S. meliloti* 1021 were performed with antibiotic resistances to gentamycin, streptomycin and tetracycline.

Root Exudates and Macerate Production. Seedlings of *Medicago sativa* cv. Marina were used to produce root exudates and macerates. *M. sativa* seeds were sterilized in ethanol 99% for 1 minute, repeatedly washed in sterilized water, and sterilized with three washes in NaClO 5%. After three water washes, seeds germinated in sterile plastic Petri dishes for 72 h in the dark and 48 h in the light in growth chamber with 2–3 ml of sterile ddH_2O. All the plantlets were then grown for 6–7 days, and then transferred over a filter of Whatman paper, in 50 ml tubes with 7 ml of sterile ddH_2O. Each tube contained five plantlets. After 14 days of incubation in the growth chamber, plantlets were removed and the exudates were extracted, filtered and stored at −20°C. Plantlets were used to produce root macerate by removing the stem, adding 5 ml of ddH_2O, and crushing the roots of 25 plantlets with mortar and pestle. One recovered, root macerate was centrifuged, filtered, and stored at -20°C.

Induction assays. All strains of *S. meliloti* 1021 carrying recombinant pOT2 plasmid were grown for 72 h at 30°C in 100 ml flasks using M9 and root exudate or root macerate as inducers. Optical density was measured using a spectrophotometer, whereas levels of GFP expression were quantified using a fluorescence microplate reader (Tecan Infinite 200 PRO, Tecan, Switzerland). As recent studies demonstrated, pSym conjugal transfer is derepressed via RctB. Indeed, *rctB* gene is required to antagonize repression by RctA[5]. In fact, the next induction test protocol was modified and applied to selected recombinant strains of *S. meliloti* 1021 carrying both the promoter:pOT2 and the pTE3rctB plasmids. These strains were used as positive controls in subsequent induction tests, which were conducted in triplicate in 10 ml Petri dishes with TY medium, in presence of root exudate or macerate, and in presence or not of a recipient strain of *S. meliloti* 1021 lacking the megaplasmid pSymA. To be able to detect every single cell fluorescence emission, a new experimental protocol was developed in which levels of fluorescence of GFP were evaluated by flow cytometry considering one million of events. Before flowing cells into the cell flow cytometer, the inducted colonies were recovered from the plate and suspended in 1 ml of physiological solution (0,9% NaCl). Cell suspensions were centrifuged and then fixed with 100% methanol for 15 minutes. After the incubation, cells were washed three times with physiological solution and, finally, resuspended in 1 ml of physiological solution.

Results

1. *Construction of derivative S. meliloti strains*

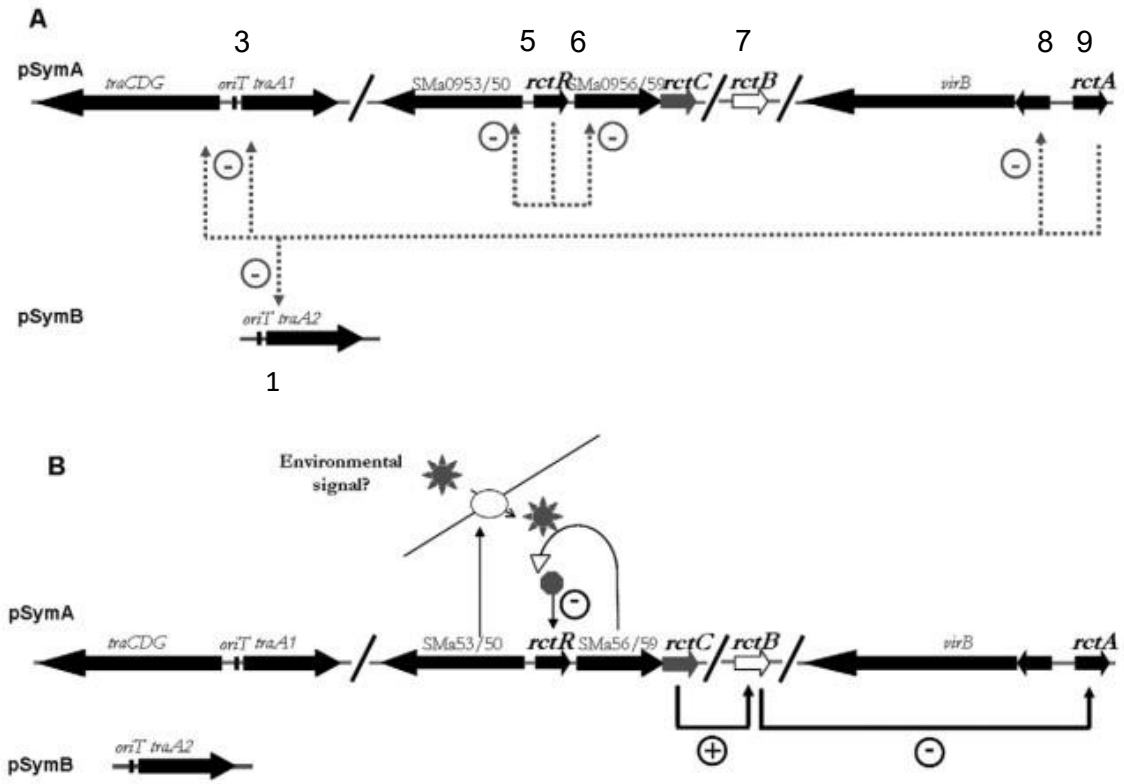


Figure 1. *Proposed model for regulation of conjugal transfer of the S. meliloti symbiotic plasmids. Numbers refer to the pairs of primers listed in Table 2. A. Under non-favourable conditions (i.e. standard laboratory media), RctR silences the SMa0953/50 and SMa0956-58-59-rctC operons, so rctB remains inactive and RctA represses conjugal transfer genes in both pSyms. B. Conjugation is achieved after an unknown compound, putatively imported by the ABC transporter encoded by SMa0953/50 and modified by the SMa0956-58-59 gene products, binds RctR thereby relieving transcription of rctC. This activates rctB to counteract RctA, conjugal transfer operons are expressed, and plasmid conjugation can proceed. +, activation; -, repression or inhibition; BM, bacterial membrane*[5]*.*

Rct system is a regulatory cascade that controls the plasmids conjugal transfer in *S. meliloti*[5], as reported in Figure 1. The molecular mechanisms which allow the activation of the regulatory cascade still unknown. In this study, we selected promoters of seven key genes belonging to the rct system to test their expression in presence of specific inducers. First of all, we amplified through PCR the promoters' sequences with specific primers (reported in Table 2) and then we cloned them into the promoterless cloning vector pOT2 through the use of two restriction enzymes system. Each recombinant pOT2 plasmid was transferred into *Escherichia coli* DH5α and then mobilized into *S. meliloti* 1021 strain. A second plasmid, the pTE3rctB plasmid, was introduced into recombinant *S. meliloti* strains. In the following Table 3, all recombinant *S. meliloti* strains we created in this experiment are described.

Table 3. *List of recombinant S. meliloti 1021 strains we created and tested in this experiment.*

Code	Strain	Plasmids	Antibodies resistances
LC01	*Sinorhizobium meliloti* 1021	traA2:pOT2	Str^R; Gm^R
LC02	*Sinorhizobium meliloti* 1021	traA1:pOT2	Str^R; Gm^R
LC03	*Sinorhizobium meliloti* 1021	rctR:pOT2	Str^R; Gm^R
LC04	*Sinorhizobium meliloti* 1021	rctC:pOT2	Str^R; Gm^R
LC05	*Sinorhizobium meliloti* 1021	rctB:pOT2	Str^R; Gm^R
LC06	*Sinorhizobium meliloti* 1021	virB:pOT2	Str^R; Gm^R
LC07	*Sinorhizobium meliloti* 1021	rctA:pOT2	Str^R; Gm^R
LC09	*Sinorhizobium meliloti* 1021	traA2:pOT2, pTE3rctB	Str^R; Gm^R, Tet^R

LC11	*Sinorhizobium meliloti* 1021	traA1:pOT2, pTE3rctB	Str^R; Gm^R, Tet^R
LC13	*Sinorhizobium meliloti* 1021	rctR:pOT2, pTE3rctB	Str^R; Gm^R, Tet^R
LC15	*Sinorhizobium meliloti* 1021	Sma0956-rctC:pOT2, pTE3rctB	Str^R; Gm^R, Tet^R
LC17	*Sinorhizobium meliloti* 1021	rctB:pOT2, pTE3rctB	Str^R; Gm^R, Tet^R
LC19	*Sinorhizobium meliloti* 1021	virB:pOT2, pTE3rctB	Str^R; Gm^R, Tet^R
LC21	*Sinorhizobium meliloti* 1021	rctA:pOT2, pTE3rctB	Str^R; Gm^R, Tet^R

2. *Evaluation of overall expression*

First of all, a first induction test was conducted on all *S. meliloti* 1021 strains carrying promoter:pOT2 plasmids using the presence of root exudate and plant macerate as conditions to be tested. Optical density and fluorescence level were measured and normalised to the negative control (*S. meliloti* 1021 harbouring the promoterless pOT2 plasmid). As shown in Figure 2a, no differences among strains were observed, except for the recombinant strain containing the Sma0956-rctC:pOT2. Comparing these results with preliminary induction tests, the expression of *rctC* gene seems to be constitutively overexpressed.
These data highlighted that the expression of single promoters belonging to the rct regulatory cascade cannot be detected through the measurement of the fluorescence levels emitted by each strain culture through the spectrophotometer. Indeed, previous studies demonstrated that the pSym plasmids transfer is a low efficiency and low frequency process[5,9–11], and we expected that only a limited number of cells would activate the promoters of rct regulatory system. Moreover, little is known about the external and internal molecular signals which carry to the activation of the rct regulatory cascade.

3. *Setting up of cytofluorimetric assay for counting expression events*

To detect and highlight the rare events in which a restricted fraction of cells expresses the GFP protein signal, in the following experiment in which every recombinant strain was inducted in presence of plant macerate and in presence of a recipient *S. meliloti* 1021 strain lacking the pSymA megaplasmid, the fluorescence emission was measured through a cytofluorimeter on one million of events. Before the cytofluorimeter reading, induced colonies were fixed in methanol for 15 minutes. After washes in physiological solution, cells were flowed into the instrument. Results showed that no significant differences in GFP expression was detected comparing the positive events counted for each *S. meliloti* recombinant strain carrying the promoter:pOT2 plasmid. These data confirmed the previous theory according to which the activation of rct conjugal transfer is a rare and stochastic event.
Based on these results, further induction assays were performed using the same protocol and the plant macerate as main inducer. In particular, we focused our attention on the comparison between two *S. meliloti* 1021 stains, those carrying the virB:pOT2 and the rctA:pOT2 recombinant plasmids. In Figure 2b, we reported the results obtained by calculating the mean of absolute frequencies of positive events for each strain in relation to one million of total events. Comparing the two induced strains, levels of GFP expression were higher in the *S. meliloti* 1021 strain which overexpresses *rctA* gene, but no significant differences were found when *rctB* gene is overexpressed in both strains.

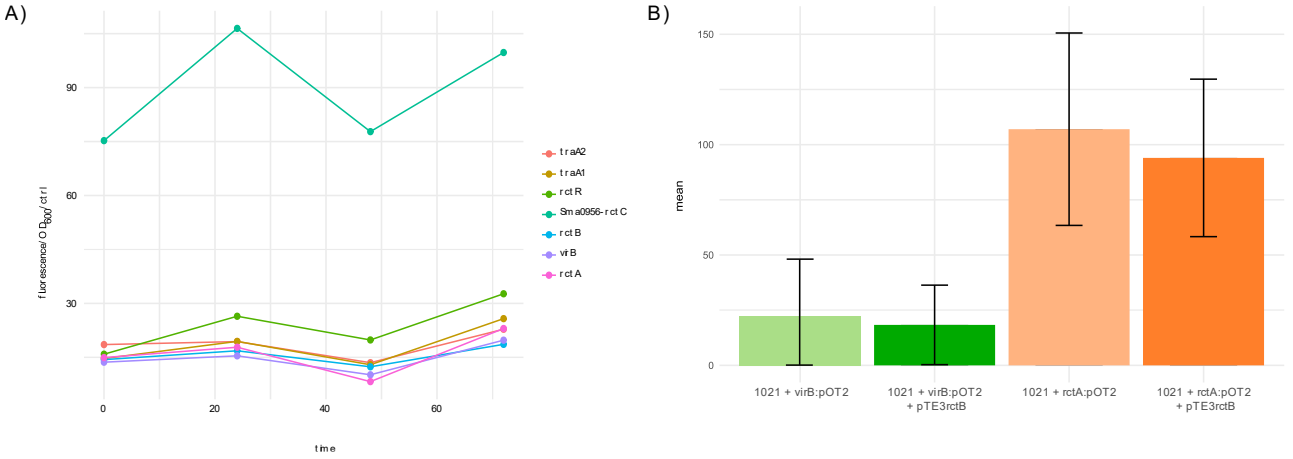


Figure 2. *GFP level expression. Panel A) shows the fluorescence emitted by S. meliloti 1021 strains carrying the recombinant promoter:pOT2 plasmid measured at 0, 24, 48 and 72 hours after the induction. GFP fluorescence was normalized in relation to the optical density and the negative control (S. meliloti 1021 with pOT2 plasmid); panel B) shows the mean and the standard deviation of the positive events of S. meliloti 1021 strains carrying virB:pOT2 or rctA:pOT2 plasmids measured through the cytofluorimeter in absence and in presence of pTE3rctB plasmid after 72 hours of induction.*

Conclusions

To date, recent studies demonstrated that conjugal transfer responds positively to population density[12] and RctA-regulated plasmids which would only transfer under conditions that disturb the activity of the transcriptional repressor RctA[5]. The general aim of this study was to understand which are the conditions that allow to repress the *rctA* expression, leading to the activation of the rct system and hence conjugal transfer of the symbiotic megaplasmid pSymA. Identification of such conditions would greatly increase our understanding of rhizobia population genomics and genome evolution, and in general of the importance of environmental conditions on horizontal gene transfer. Moreveor, the possibility to reproduce such conditions in laboratory scale may open novel opportunities for genome manipulation in bacteria and synthetic biology, including production of hybrid rhizobia strains carrying different arrangements of symbiotic megaplasmids.

Here, we have set up a panel of gene-reported systems for the scanning of the rct regulatory cascade and defined protocols for testing expression in plate reader and cytofluorimeter. Preliminary results on the recombinant strains confirmed that that under laboratory conditions the conjugal transfer of megaplasmid pSymA is highly repressed. However, a small fraction of cells expressing the conjugal transfer machinery was detected, though inducing conditions have been not identified yet.

Bibliography

1. Meade, H. M., Long, S. R., Ruvkun, G. B., Brown, S. E. & Ausubel, F. M. Physical and genetic characterization of symbiotic and auxotrophic mutants of Rhizobium meliloti induced by transposon Tn5 mutagenesis. *J. Bacteriol.* **149**, 114–122 (1982).
2. Grant, S. G., Jessee, J., Bloom, F. R. & Hanahan, D. Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. *Proc. Natl. Acad. Sci.* **87**, 4645–4649 (1990).
3. Finan, T. M., Kunkel, B., De Vos, G. F. & Signer, E. R. Second symbiotic megaplasmid in Rhizobium meliloti carrying exopolysaccharide and thiamine synthesis genes. *J. Bacteriol.* **167**, 66–72 (1986).
4. Karunakaran, R., Mauchline, T. H., Hosie, A. H. F. & Poole, P. S. A family of promoter probe vectors incorporating autofluorescent and chromogenic reporter proteins for studying gene expression in Gram-negative bacteria. *Microbiology* **151**, 3249–3256 (2005).

5. Nogales, J., Blanca-Ordóñez, H., Olivares, J. & Sanjuán, J. Conjugal transfer of the *Sinorhizobium meliloti* 1021 symbiotic plasmid is governed through the concerted action of one- and two-component signal transduction regulators: Regulation of *S. meliloti* pSym conjugal transfer. *Environ. Microbiol.* **15**, 811–821 (2013).
6. diCenzo, G. C., MacLean, A. M., Milunovic, B., Golding, G. B. & Finan, T. M. Examination of Prokaryotic Multipartite Genome Evolution through Experimental Genome Reduction. *PLoS Genet.* **10**, e1004742 (2014).
7. Dd Dienst, D. & Eriksson, H. *Preparation of Chemical Competent E. Coli Cells (w/ Calcium Chloride) V1*. https://www.protocols.io/view/preparation-of-chemical-competent-e-coli-cells-w-c-qe4dtgw (2018) doi:10.17504/protocols.io.qe4dtgw.
8. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual.* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).
9. Galardini, M., Pini, F., Bazzicalupo, M., Biondi, E. G. & Mengoni, A. Replicon-dependent bacterial genome evolution: the case of Sinorhizobium meliloti. *Genome Biol. Evol.* **5**, 542–558 (2013).
10. Achchi, H. G. & Hemananda, K. M. Genomic and Proteomic Approaches to Understanding Conjugation in Rhizobium leguminosarum VF39. (2023).
11. Blanca-Ordóñez, H. *et al.* pSymA-Dependent Mobilization of the *Sinorhizobium meliloti* pSymB Megaplasmid. *J. Bacteriol.* **192**, 6309–6312 (2010).
12. Danino, V. E., Wilkinson, A., Edwards, A. & Downie, J. A. Recipient-induced transfer of the symbiotic plasmid pRL1JI in Rhizobium leguminosarum bv. viciae is regulated by a quorum-sensing relay. *Mol. Microbiol.* **50**, 511–525 (2003).

The impact of synthetic microbial consortia, fertilization regimes and maize growth stages on taxonomic and functional microbiome under temperate climate conditions

Authors: A. Bevivino[1*], L. Cangioli[2], N. Kleinbölting[3,4], A. Fiore[1], A. Schlüter[3], A. Sczyrba[3], A. Mengoni[2], J. Hett[5], M. Costanzo[1], F. Sevi[1], G. Aprea[1], S. Tabacchioni[1], A. Pihlanto[6], D. Neuhoff[5]

[1] ENEA, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Department for Sustainability, Rome, Italy
[2] Center for Biotechnology (CeBiTec), Bielefeld University, 33615 Bielefeld, Germany
[3] Forschungszentrum Jülich, Jülich, Germany
[4] Department of Biology University of Florence, Florence, Italy
[5] Institute of Crop Science and Resource Conservation, Dep. of Agroecology & Organic Farming, University of Bonn, Bonn, Germany
[6] Natural Resources Institute Finland (Luke), Helsinki, Finland

Abstract

The growth of crop plants relied heavily on chemical fertilizers and pesticides with serious environmental consequences and negative effects on human health. Environmentally friendly solution to develop more sustainable agricultural practices to offset the damage caused by many agrochemicals are necessary[1]. Efforts are being made in developing new inoculants microbial consortia not only for growing crops sustainably but also for preserving soil health[2]. Additionally, the question of whether natural soil microbiomes are negatively affected by adding foreign microorganisms must also be addressed. It is also of great interest to investigate the impact that PGPM inoculations have on indigenous soil microbiomes by evaluating the microbial baseline of soil prior to inoculation as well as after. Previous studies that have measured the non-target effects of PGPM on the indigenous bacterial community have demonstrated contrasting results, ranging from a generally rare, transient, and spatially limited significant effect to a direct effect on the plant microbiome due to bacterial inoculation[3,4]. Here, we applied 16S rRNA gene amplicon sequencing to examine the effects of microbial consortia inoculation on the non-target rhizosphere soil microbiome at different plant growth stages and fertilization levels. In this context, non-target effects of inoculation with MC_B and MC_C and Microsat consortia on rhizosphere soil microbiome were evaluated. The diversity, composition, and relative abundance of bacterial community in rhizosphere soil were represented at different sampling times during the maize growth, via Illumina high-throughput sequencing.

Introduction

The soil is a highly complex terrestrial environment and represents the hotspot of a myriad of agriculturally beneficial microorganisms. These beneficial microorganisms interact both with one another and with plants and can improve soil quality and enhance crop productivity. Plant-growth promoting bacteria have direct and indirect effects on both crop growth and quality, and therefore have positive effects on agricultural sustainability as they contribute to the resistance and resilience of agroecosystems to abiotic disturbance and stress[5]. Microbial inoculants have the potential to be used as “plant probiotics” for a sustainable agriculture because of their multiple roles such as nutrient supplementation, plant growth promotion, soil health enhancement, disease suppression by providing resilience, especially under harsh environmental conditions. Moreover, unlike fertilizers, biofertilizers do not have any harmful effects on the ecosystem[6]. Different microbial inoculants are used individually or in the combination of different types of microorganisms as consortia (SynComs) for evergreen agriculture[7]. However, knowledge of their effect on

resident microbial communities and interactions with the indigenous soil microbes is scant and controversial, ranging from a generally rare, transient, and spatially limited significant effect to a direct effect on the plant microbiome[6]. According to available literature, the extent of microbiome modulation depends on the on the time of sampling (days or weeks after inoculation)[8]. Many studies revealed shifts in soil microbiome composition directly after bioinoculant application. A few weeks after the application, shifts are often not evident anymore. In general, six different types of microbial inoculant impact on the indigenous soil-plant microbiome were described pointing out the following types of microbiome modulations: (i) transient microbiome shifts, (ii) stabilization or increase of microbial diversity, (iii) stabilization or increase of microbiome evenness, (iv) restoration of a dysbiosis/compensation or reduction of a pathogen-induced shift, (v) targeted shifts toward potential beneficial phyla of the indigenous microbiota, and (vi) suppression of potential pathogens[8].

Beneficial microbial inoculants are a complementary management tool to improve soil fertility and increase vegetable production. However, its integration with chemical fertilization has been poorly studied. The application of chemical fertilizer can indirectly impact soil microorganisms by changing soil properties or directly by input of nutrients[9,10]. Understanding the changes in soil microbial communities under different fertilization regimes is needed for determining the optimal fertilization methods to improve soil conditions and crop productivity. Over the past decade, microbiome studies have offered exceptional opportunities to understand the microbial community structure in agricultural sciences with diverse aims, including diversity profiling and functional characterization[11]. This approach is important, since microbial community structure provide insights into the characteristics of microbes as determined by functional gene composition including the microbial composition (i.e., the identity and relative abundances of all taxa)[12]. Analyses of metagenomic data typically seek to identify the specific organisms, operational taxonomic units, or pathways whose relative abundances differ between two or more groups of samples. As in clinical settings metagenomic biomarkers have been proposed as potential biomarkers for various disease states[13], in agriculture several features of microbial communities can be used as biomarkers of soil fertility and crop productivity.

The main objectives of the present work are to assess the impact of specifically selected microbial consortia (MC) under different fertilization regimes on taxonomic and functional rhizosphere microbiome on maize under temperate conditions. MC were applied alone or in combination with Arbuscular Mycorrhizal Fungi (AMF) in open field under organic management and compared with commercial microbial products. The effect of microbial consortia inoculation on the non-target rhizosphere soil microbiome was investigated by a coupled approach based on 16S rRNA gene sequencing and shotgun metagenomic sequencing and bioinformatic analysis.

Methods

Experimental fields, Sampling and Samples Treatment

In Germany, the field trial was conducted at the organic research farm “Wisengut” (50.480 N, 7.170 E), near to Bonn, on Maize cv. Benedictio with a two factorial split plot design. The main plot refers to the PGPMs treatment with the four factor levels i) control, ii) MC_B, iii) MC_C and iv) MICOSAT F and the subplot potato protein liquid (PPL) fertilization with the three factor levels 50 kg, 100kg, 200 kg mineral nitrogen per ha. In total 48 plots were included consisting of 3 x 4 = 12 treatments repeated four times. The inoculation of the seeds with the microbial treatment (MT) was carried out in the form of a seed coating one day before sowing. The MT, stabilized in zeolite, were used for coating at a rate of 0.2 g per grain thanks to a starch binder. Twelve weeks after sowing, plots were re-inoculated with MT treatments dispersed in water. Within each treatment, rhizosphere soil was collected from 6 individual plants as follows. The entire root was taken out from the field paying attention to the integrity of the root wearing gloves. The non-adhering soil was

manually shaken from the roots. Then, five-six grams of fresh rhizosphere soil were collected by detaching soil with a sterile scalpel from different parts of the plant root system. Fresh rhizosphere soil was kept in a zipped plastic bag and stored at -20°C until use.

DNA extraction from rhizosphere soil samples

DNA extraction from rhizospheric soil samples was performed using QIAGEN's new DNeasy PowerSoil Pro Kit (QIAGEN Group, Hilden, Germany) according to the manufacturer's instruction. To measure the integrity of samples, DNAs were run on 1.5% agarose gel electrophoresis with 1X TAE buffer (Sigma Aldrich), at 100 V for 45 min, and subsequently stained with GelRed (0.5µl mL-1), visualized and photo-documented under ultraviolet light on Bio Red Molecular Imager ChemiDoc TM XRS+, US. The evaluation of the quantity and quality of DNAs was done by measuring DNA extracted samples using NanoDrop 2000 (Thermo Fisher Scientific) and Qubit 4.0 fluorometer (Invitrogen, Life technologies).

In addition, DNA quality was checked by PCR amplification using 16S rRNA gene primers pairs 314F and 805R that amplify a 500 bp amplicon of V3-V4 region of 16S rDNA gene. PCR conditions: initial incubation for 1 min at 94°C followed by 25 cycles of 30 s at 94°C, 30 s at 55°C, 45 s at 72°C with a final incubation for 7 min at 72°C. For each sample, two independent DNA extractions were performed and then pooled them in an equimolar ratio, obtaining a composite associated DNA sample for each plant. Samples were stored at -20° C prior to use.

Library preparation and sequencing

After checking the quality of PCR product through 1.5% agarose gel electrophoresis, Illumina 16S library construction and metagenomic sequencing were performed using Illumina MiSeq platform and 300 PairedEnds strategy at MATIS facilities (Reykjavík, Iceland). Briefly, V3-V4 regions of 16S rRNA gene were amplified using the following primers: Pro341F, 5'- CCTACGGGNBGCASCAG-3', and Pro805R, 5'-GACTACNVGGGTATCTAATCC-3' (Takahashi et al. 2014). Primers were modified with forward and reverse overhangs (5'- TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-[locus-specific sequence]-3' and 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-[locus-specific sequence]-3', respectively) necessary for dual index library preparation following Illumina MiSeq protocol (https://web.uri.edu/gsc/files/16s-metagenomic-library-prep-guide-15044223-b.pdf.) Samples were normalized, pooled and run on the Illumina MiSeq with 2x300 bp approach.

Database creation, pre-processing of raw reads and taxonomic assignment

The database of the project's data is implemented as data frame object in R. Briefly, data stored csv-delimited file was imported in RStudio environment version 4.0 using the R function *read.csv()*. Raw reads were previously demultiplexed and adapters sequences were trimmed using Trim Galore! (version 0.6.3) on Galaxy server (link: https://usegalaxy.org/). Further data analysis is carried out on the RStudio software. Paired end Illumina MiSeq data were processed following the DADA2 pipeline using the '*dada2*' R-package v. 1.16.0[14]. This pipeline corrects errors that are introduced into sequencing data during amplicon sequencing, and it includes steps to filter, dereplicate, identify chimeras, and merge paired-end reads. Samples with lower than 5000 reads were discarded. Taxonomic assignment to the sequence variants and classification was performed using '*DECIPHER*' R package[2] (version 2.0.2)[15], a toolset for deciphering and managing biological sequences. DADA2 pipeline outputs a table which contains the amplicon sequence variants (ASVs), that are of higher resolution than OTUs.

Differential abundance and multivariate analysis

A phyloseq object was created with the '*phyloseq'* R package (version 1.34.0)[16] and '*vegan*' R package (version 2.5-7)[17] to perform differential abundance analyses. 'ggplot2' R package (version 3.3.5)[18], was carried out to create graphs. PerMANOVA tests were performed using the '*pairwise.Adonis*' R package (version 0.0.1)[19]. Alpha diversity, beta diversity and relative abundance of keystone taxa studies were carried out using the following R functions: plot_richness(), plot_bar(), estimate_richness(), pairwise.wilcox.test(), adonis(), pairwise.adonis(), plot_ordination(). Good's coverage was calculated using the '*QsRutils'* R package (version 0.1.5)[20] *goods()* function.

A cantered log-ratio transformation was applied to the samples using the '*microbiome'* R package (version 1.12.0)[21]. Differential ASVs counts analysis was performed using the '*DESeq2'* R package (version 1.30.0) functions: phyloseq_to_deseq2(), estimateSizeFactors(), DESeq(), results() (Love et al. 2017). A PCA analysis was carried out after a vst transformation using the following functions: varianceStabilizingTransformation(), plotPCA(), ASVs.

Results

Taxonomic composition and alfa-diversity of maize rhizosphere microbiota

16S rRNA gene amplicon sequencing was performed using the NGS sequencer MiSeq by Illumina and for most of samples a satisfactory number of reads was reached. Goods' coverage values of selected samples reached above 0.99.

Alfa diversity of maize microbiota was performed for each group of samples by measuring the Shannon, Simpson, Chao1, Fisher, ACE (abundance-based coverage estimators) and Richness indices (data not shown). In particular, Wilcoxon test on Richness index shows that sampling time has no influence on diversity, while treatment and fertilization do. (Table 1).

Table 1. *p.values obtained performing separate Wilcoxon tests maize samples among treatments, levels of fertilization and time of sampling, respectively. Significant p.values (p.value<0.05) are reported in red.*

Treatment	**MC_B**	**MC_C**	**Control**
MC_C	0.726	-	-
Control	0.726	0.317	-
Micosat	0.212	0.726	0.019

Fertilization	**100 kg N/ha**	**200 kg N/ha**
200 kg N/ha	0.022	-
50 kg N/ha	0.236	0.111

Time of sampling	**1 week**	**Flowering**
Flowering	0.36	-
Harvesting	0.49	0.23

Concerning the taxonomic distribution of the amplicon sequence variations (ASVs), those associated with Actinobacteriota, Proteobacteriota, Firmicutes, Verrucomicrobiota, Chloroflexi and Acidobacteriota were found to be the most abundant (Figure 1).

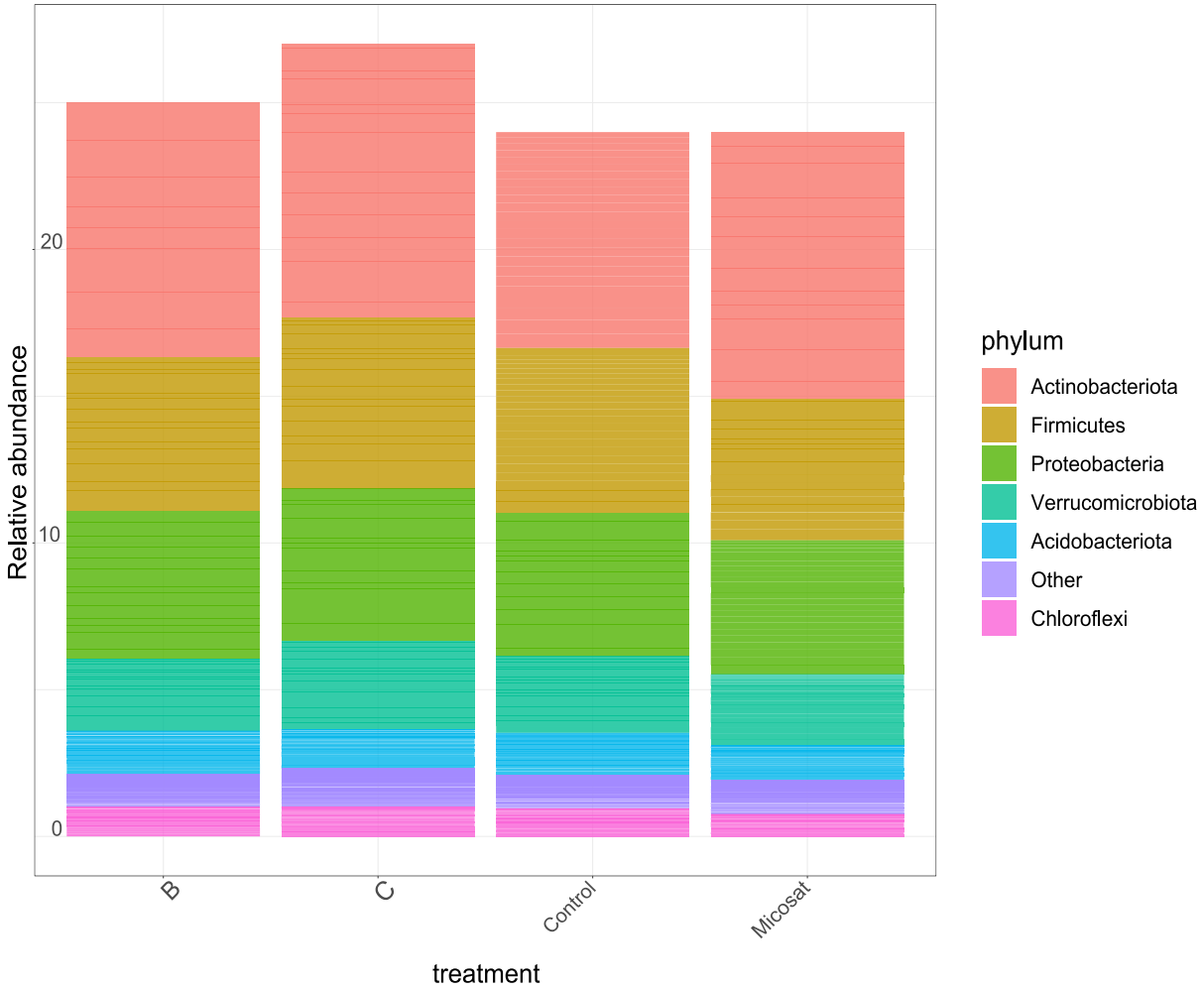


Figure 1. *Taxonomic composition of maize rhizosphere microbiota from field trials in Germany.*

Analysis of beta-diversity of rhizosphere microbiota

A nonMetric Multidimensional Scaling was then used to identify clusters of samples with homogeneous taxonomic composition. A possible patterning related to inoculum x treatment interaction is present (Figure 2), while for each group of samples, no differences were found among treatments and time of sampling.

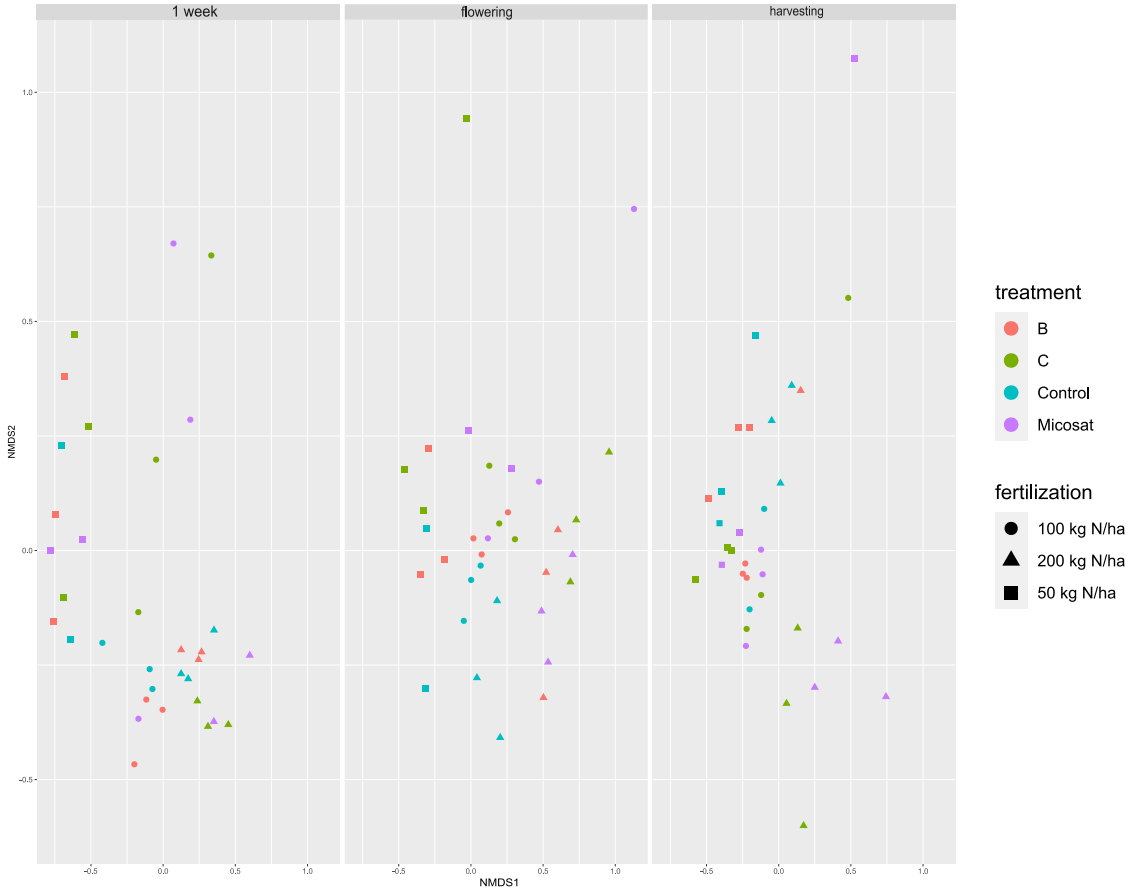


Figure 2. *nonMetric Multidimensional Scaling on maize samples, divided by sampling time.*

To inspect if statistically significant differences in microbiota composition between treatments and time point were present, permutational multivariate analysis of variance (PERMANOVA) was performed. A significant effect of sampling times was found ($p<0.05$) indicating that the taxonomic composition changed over time (Table 2), as expected due to ecological succession of communities possibly related to plant growth stages (see for instance Cangioli et al., 2022[4]). No interaction effect between consortia MC_B and MC_C and control was found ($p>0.05$) meaning that the application of such microbial consortia had no effect over the taxonomic composition of maize rhizosphere microbiota; in contrast, an interaction between consortium Micosat and both control and consortium MC_B was found ($p<0.05$). The community composition in the higher fertilizer treatment (200 Kg N/ha) was different from those in lower (50 kg N/ha) and medium (100 kg N/ha) synthetic fertilizer treatments.

Table 2. *p.values obtained performing PERMANOVA on centered log-ratio transformed samples among treatments, fertilization and time of sampling, respectively. Significant p.values (p.value<0.05) are reported in red.*

Treatment	**B**	**C**	**Control**	**Micosat**
B	-	0.656	0.932	0.043
C	0.625797	-	0.587	0.116
Control	0.936994	0.583113	-	0.041
Micosat	0.03919	0.108457	0.037874	

Fertilization	**100 kg N/ha**	**200 kg N/ha**	**50 kg N/ha**
100 kg N/ha	-	0.002	0.61
200 kg N/ha	0.00098317	-	0.002
50 kg N/ha	0.64440695	0.00334867	-

Time of sampling	**1 week**	**Flowering**	**Harvesting**
1 week	-	1.00E-03	0.016
flowering	8.89E-05	-	0.032
harvesting	1.15E-02	2.64E-02	-

To further investigate taxonomic differences among samples, differences in ASVs abundances among samples were considered and used to perform differential abundance analysis (DESeq2). Results from this analysis were then further ordered using a PCA with the type of treatment, fertilization, and time of sampling, as variables (Figure 3).

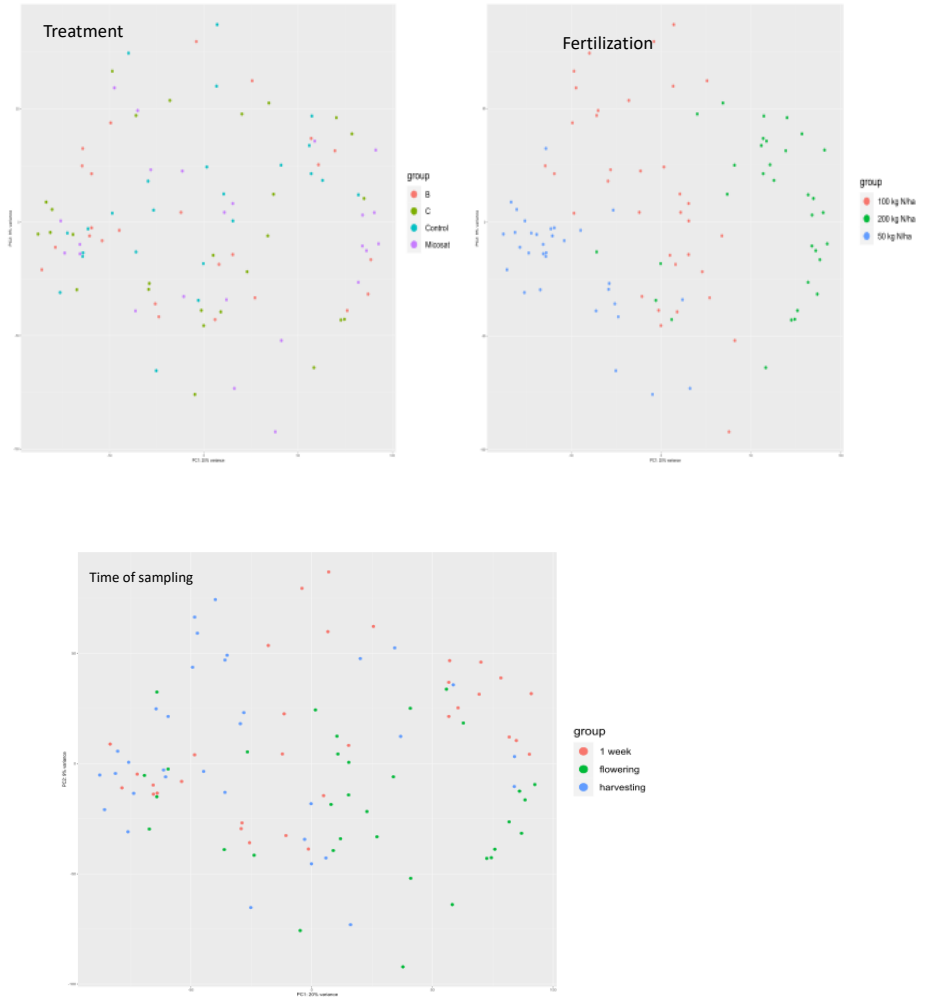


***Figure 3**. PCA on DeSeq2 data (differential abundance) referring to treatment with consortia, fertilization level and sampling time (data from the first year of experiment).*

No evident grouping regarding the treatments was found, whereas three evident groups corresponding to the different types of fertilization were detected (Figure 4), suggesting that ASVs in the rhizosphere microbial community structures varied in relation to the quantity of fertilizer supplied, as indicated also in previous PERMANOVA.

Discussion

The heavy use of chemical fertilizers and pesticides in modern agriculture has negatively impacted soil and human health. The urgent need to develop new sustainable and ecofriendly solutions represents an increasing topic that attracts the attention of many researchers[22].

Soil microbiome engineering offers an effective alternative to the usage of chemical products in agriculture and, in association with this, the development of synthetic microbial communities to apply directly as bioinoculants in the soil is becoming an essential strategy to increase crops yield and to promote plant growth[7]. The aim of this work was to investigate the impact that microbial consortia applications could have on the taxonomic composition of native soil microbial community of maize plants in relation to different plant growth stages and fertilization levels through the sequencing of the 16S rRNA gene.

In contrast to previous studies in which plant phenological stage plays an important role in rhizomicrobiota diversity and composition[4,23], here we demonstrated that alpha and beta diversity of maize microbiota are

influenced by the different microbial consortia application and the fertilization level, indicating a rhizosphere microbiota patterning related to inoculum x treatment interaction is present.
On the other hand, a significant effect of time of sampling was found in microbiota composition through a permutational multivariate analysis of variance (PERMANOVA), indicating that the taxonomic composition changed over the time, probably in associations with different plant growth stages[4].
Nowadays, many microbial consortia have been commercialized and applied to promote plant resistance, growth and yield[3,10,24], but one of the most important aspect to be considered after the inoculation of biofertilizers is their impact on native soil communities. Our results suggested that no effect on the structure of microbiota of MC_B and MC_C with respect to the control was found, meaning that the application of SIMBA microbial consortia was neutral with respect to the native maize rhizosphere microbiota; in contrast, a significant difference between Micosat and both control and MC_B was found, indicating that Micosat exert a disturbance effect over the microbiota.
According with Cangioli et al., 2022[6], when considering different levels of fertilization, we observed differences in rhizomicrobiota composition. In fact, it was affected by the high fertilizer treatment (200 Kg N/ha) and it was different from those in low (50 kg N/ha) and medium (100 kg N/ha) synthetic fertilizer treatments. This evidence on soil microbial community diversity and taxonomic composition may be caused by direct effects of nitrogen enrichment as a nutrient, or by indirect changes in soil and plant properties, as described in Zeng et al., 2016[25].
Overall, the microbiota analyses of rhizospheric soil showed the resilience of native rhizosphere microbial communities in response to SIMBA's microbial consortia (SMC), suggesting that the inoculation with microbial consortia causes no level of disturbance and changes over the composition and diversity of soil microbial community. Therefore, microbial consortia MC_B and MC_C represent good candidates as plant probiotics for a sustainable agriculture.
Nevertheless, we are aware that this study has some limitations due to the duration of only one year of sampling and further studies are required to investigate the effectiveness of long-term impacts of microbial consortia inoculation on soil microbiota.

Bibliography

1. Arif, I., Batool, M. & Schenk, P. M. Plant Microbiome Engineering: Expected Benefits for Improved Crop Growth and Resilience. *Trends Biotechnol.* **38**, 1385–1396 (2020).
2. Afridi, M. S. *et al.* Harnessing microbial multitrophic interactions for rhizosphere microbiome engineering. *Microbiol. Res.* **265**, 127199 (2022).
3. Bianco, C. *et al.* Endophytes from African Rice (Oryza glaberrima L.) Efficiently Colonize Asian Rice (Oryza sativa L.) Stimulating the Activity of Its Antioxidant Enzymes and Increasing the Content of Nitrogen, Carbon, and Chlorophyll. *Microorganisms* **9**, 1714 (2021).
4. Cangioli, L. *et al.* Effect of Site and Phenological Status on the Potato Bacterial Rhizomicrobiota. *Microorganisms* **10**, 1743 (2022).
5. Alori, E., Dare, M. O. & Babalola, O. Microbial Inoculants for Soil Quality and Plant Health. in vol. 22 281–307 (2017).
6. Cangioli, L. *et al.* Differential Response of Wheat Rhizosphere Bacterial Community to Plant Variety and Fertilization. *Int. J. Mol. Sci.* **23**, 3616 (2022).
7. de Souza, R. S. C., Armanhi, J. S. L. & Arruda, P. From Microbiome to Traits: Designing Synthetic Microbial Communities for Improved Crop Resiliency. *Front. Plant Sci.* **11**, 1179 (2020).

8. Berg, G., Kusstatscher, P., Abdelfattah, A., Cernava, T. & Smalla, K. Microbiome Modulation—Toward a Better Understanding of Plant Microbiome Response to Microbial Inoculants. *Front. Microbiol.* **12**, 650610 (2021).
9. Pan, H. *et al.* Organic and inorganic fertilizers respectively drive bacterial and fungal community compositions in a fluvo-aquic soil in northern China. *Soil Tillage Res.* **198**, 104540 (2020).
10. Lu, J. *et al.* Nitrogen fertilizer management effects on soil nitrate leaching, grain yield and economic benefit of summer maize in Northwest China. *Agric. Water Manag.* **247**, 106739 (2021).
11. Li, X. *et al.* Occurrence, source, ecological risk, and mitigation of phthalates (PAEs) in agricultural soils and the environment: A review. *Environ. Res.* **220**, 115196 (2023).
12. Adeleke, B. *et al.* Strategies to Enhance the Use of Endophytes as Bioinoculants in Agriculture. *Horticulturae* **8**, 498 (2022).
13. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60 (2011).
14. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
15. Wright, E. S. DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics* **16**, 322 (2015).
16. McMURDIE, P. J. & Holmes, S. PHYLOSEQ: A BIOCONDUCTOR PACKAGE FOR HANDLING AND ANALYSIS OF HIGH-THROUGHPUT PHYLOGENETIC SEQUENCE DATA. in *Biocomputing 2012* 235–246 (WORLD SCIENTIFIC, Kohala Coast, Hawaii, USA, 2011). doi:10.1142/9789814366496_0023.
17. Oksanen, J. *et al.* Package 'vegan'. *Community Ecol. Package Version* **2**, 1–295 (2013).
18. Wickham, H. ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* **3**, 180–185 (2011).
19. Martinez Arbizu, P. pairwiseAdonis: Pairwise multilevel comparison using adonis. R package version 0.4. (2020).
20. Quensen, J. QsRutils: R Functions Useful for Community Ecology. (2020).
21. Lahti, L. & Shetty, S. *Introduction to the Microbiome R Package*. (2018).
22. Bano, S., Wu, X. & Zhang, X. Towards sustainable agriculture: rhizosphere microbiome engineering. *Appl. Microbiol. Biotechnol.* **105**, 7141–7160 (2021).
23. Chaparro, J. M., Badri, D. V. & Vivanco, J. M. Rhizosphere microbiome assemblage is affected by plant development. *ISME J.* **8**, 790–803 (2014).
24. Minchev, Z., Kostenko, O., Soler, R. & Pozo, M. J. Microbial Consortia for Effective Biocontrol of Root and Foliar Diseases in Tomato. *Front. Plant Sci.* **12**, 756368 (2021).
25. Zeng, J. *et al.* Nitrogen fertilization directly affects soil bacterial diversity and indirectly affects bacterial community composition. *Soil Biol. Biochem.* **92**, 41–49 (2016).

Genome insights into beneficial microbial strains composing SIMBA microbial consortia applied as biofertilizers for maize, wheat and tomato

L. Cangioli[1], S. Tabacchioni[2], A. Fiore[2], G. Aprea[2], P. Ambrosino[3], E. Ercole[4], S. Sorensen[5], A. Mengoni[1] and A. Bevivino[2*]

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

[2]Department for Sustainability, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, ENEA Casaccia Research Center, Via Anguillarese 301, 00123 Rome, Italy

[3]AGRIGES srl, San Salvatore Telesino, Italy

[4]Centro Colture Sperimentali, CCS-AOSTA srl, Quart, Italy

[5]Department of Biology, University of Copenhagen, Copenhagen, Denmark

*Author to whom correspondence should be addressed

Abstract

The genome sequences of six strains composing beneficial microbial consortia with plant growth-promoting (PGP) activities in maize, tomato and wheat have been sequenced. The genome size of strains ranged from 4.5 to 7.5 Mbp, carrying a number of genes spanning from 4402 to 6697, and a GC content of ???% to ???%. The annotation of genomes revealed a few genes implicated in functions related to antagonism, pathogenesis and other secondary metabolites possibly involved in plant growth promotion (i.e. siderophore) and gene clusters for protection against oxidative damage. Overall, the WGS of strains identified relevant functions involved in PGP and biocontrol activities as antibiotic synthesis gene clusters and siderophore production. No genes putatively involved in pathogenesis were identified.

Introduction

Plant growth-promoting (PGP) bacteria are a taxonomically heterogenous group of bacteria which are able to contribute to plant growth and protection [1]. They are normally found within the plant-associated microbiota, as members of the rhizosphere or of the phyllosphere (the habitats surrounding roots and leaves, respectively) or as endophytes of different plant organs. These bacteria can contribute to the host plant health and yield with multiple abilities from the absorption of nutrients to the regulation of host immune response. They provide protection against pathogens, additional metabolic capabilities, essential nutrients, and even phytohormone modulation.

In the last years, an ever-increasing interest in plant-associated bacteria has been manifested by researchers but also growers, consumers, and policymakers. In fact, the effectiveness of PGP bacterial inoculants in contributing a net increase in crop productivity coupled with a decreased use of agrochemicals has been clearly demonstrated [2]. Interests in scouting plant-associated microbiota and culturing novel PGP bacteria is growing [3,4]. Nowadays, one of the main concerns of the utilization of inoculants composed by a single bacterial strain is that, while they may be effective under *in vitro* testing, they usually fail in field conditions due to competition with indigenous soil and plant microbiota, as well as in relation to face environmental conditions [5,6]. Recently, innovative solutions have been proposed such as the production of microbial consortia composed by different species, which can have complementary activities with respect to the plant and which can assure enough persistence in the environment to produce reliable plant growth-promoting effects [7–9]. Such microbial consortia are sometimes referred to as synthetic communities (SynComs) [4,10], when they comprise several members trying to mimic the core functions of a healthy native microbial community [11]. In many cases, SynComs have been tested *in vitro* o in greenhouse conditions [12], providing the opportunity to model the complexity of plant-microbe interactions and of the soil and rhizosphere microbiome [4,13,14]. However, the lab-to-the-field gap is often very sharp and led to

unsuccessful attempts [15]. Moreover, due to the need of experimental fields and facilities, few studies only provide testing of SynComs under field conditions [16,17].

The whole genome sequencing (WGS) offers the highest level of bacterial strain discrimination facilitating the development of tools for traceability into the soil and more targeted risk assessment to exclude any potential risk associated with scaling-up and commercial application. Genome sequencing is now an important component of natural product research.

Under the framework of H2020 SIMBA project SynComs were developed [8] and tested on maize and wheat in greenhouse conditions, as well under field conditions in maize and tomato [18–20]. We used whole-genome sequencing and annotation to provide useful insights into the mechanisms governing plant growth promoting of some members of SIMBA SynComs, not previously sequenced, successfully used in greenhouse and field studies is reported (references). The bacterial strains of the community comprise: *Azotobacter vinelandii* DSM 2289, *Bacillus amyloliquefaciens* LMG 9814, *Bacillus licheniformis* PS141, *Bacillus* velezensis BV84, *Burkholderia ambifari*a MCI7, *Pseudomonas granadensis* A23/T3c, *Pseudomonas fluorescens* DR54, and *Rahnella aquatilis* BB23/T4d.

Materials and Methods

Bacterial strains and culture conditions

Bacterial strains were isolated from soil, rhizosphere or plant tissues. Isolation features and relevant information have been already reported [8] and are summarized in Table S1.

Genome sequences and annotation

Genomic DNA from each microbial strain was extracted with sodium dodecyl sulfate proteinase K lysis buffer, followed by a treatment with cetyltrimethylammonium bromide (CTAB) as described in Current Protocols for Molecular Biology. Quantity and purity of DNA was checked

by NanoDrop (NanoDrop Technologies, USA) and gel electrophoresis. Genomic DNA was sent to IGA Technology Services (Udine, Italy) to perform paired end sequencing (HiSeq 2000 Illumina platform with 2x150 approach). Sequences were assembled with SPAdes [21]. Annotation and COG assignment were performed by Prokka 1.4.0 [22] from the Galaxy webserver Orione (https://orione.crs4.it/) (accessed on 12/09/2021).

Genome characterization and mining

Taxonomic identification of strains from genome sequence was performed using Type (Strain) Genome Server [23]. dDDH values computed with Genome BLAST Distance Phylogeny approach. Formula *d4* from [24] was used since it is independent of genome length and is thus robust against the use of incomplete draft genomes. Secondary metabolites gene clusters have been identified by AntiSMASH [25] (accessed on 01/10/2021). The presence of Type III secretion systems was analyzed by T3Sepp [26]

Pangenome analysis with Roary [27] was used to define core and dispensable/unique gene fractions against close relatives of each strain.

Data availability

Sequences are deposited in NCBI database under BioProject xxxxx

Results

Genomic sequencing report

Genome size of the six sequenced bacterial strains of the SynCom ranged from 4.5 to 7.5 Mbp (Table 1). The number of genes spanned from 4402 to 6697 in agreement with corresponding genome sizes. Complete genomes of the remaining 2 strains, *Azotobacter vinelandii* DSM 2289 (https://genomes.atcc.org/genomes/457dab6207084f7c) and *Bacillus amyloliquefaciens* LMG 9814

(https://genomes.atcc.org/genomes/d100fd1dcdab4a05), were previously reported and contained 4'829 and 4'032 genes, respectively.

Table 1. Genome sequencing statistics and annotation.

Strains	Contigs	CDSs	Genes	Genome size (bp)	rRNA	tRNA	tmRNA	Repeat_region
***Bacillus velezensis* BV84**	9'042	5'140	5'243	6'928'355	10	92	1	/
***Bacillus licheniformis* PS141**	1242	4'318	4'402	4'504'417	6	77	1	/
***Burkholderia ambifaria* MCI7**	274	6'618	6'697	7'509'419	4	74	1	/
***Pseudomonas granadensis* A23/T3c**	1870	5'835	5'903	7'034'766	5	62	1	/
***Pseudomonas fluorescens* DR54**	359	5'420	5'501	6'178'566	4	68	1	/
***Rahnella aquatilis* BB23/T4d**	509	5'087	5'176	5'606'795	6	82	1	1

For each genome, *in silico* DNA-DNA hybridization (DDH) and genome-based phylogeny were computed to allow a robust taxonomic classification of the strains [24]. Results of taxonomic assignment based on dDDH values are reported in Table 2. Phylogenetic trees for each strain are shown in Supplementary Figure S1.

Table 2. Taxonomic classification of strains. dDDH (formula *d4* along with confidence intervals C.I.) and GC% difference with the best hit is shown.

Query strain	Subject strain	dDDH (d4, in %)	C.I. (d4, in %)	G+C content difference (in %)
Bacillusvelezensis BV84	*Bacillus velezensis NRRL B-41580*	83,8	[81.0 - 86.2]	3,3
Bacillus licheniformis PS141	*Bacillus licheniformis ATCC 14580*	97,2	[96.1 - 98.0]	0,21

Burkholderia ambifaria MCI7	*Burkholderia ambifaria AMMD*	79,2	[76.3 - 81.9]	0,04
Pseudomonas granadensis A23/T3c	*Pseudomonas crudilactis UCMA 17988T*	47,1	[44.5 - 49.7]	0,45
Pseudomonas fluorescens DR54	*Pseudomonas carnis B4-1T*	77,7	[74.7 - 80.4]	0,23
Rahnella aquatilis BB23/T4d	*Rahnella aceris SAP-19*	95,7	[94.2 - 96.8]	0,13

118

119

120 *Prediction of functions related to antagonism and plant growth promotion*

121 The presence of functions related to antagonism, pathogenesis and other secondary metabolites
122 possibly involved plant growth promotion (i.e. siderophore) was inspected on the whole dataset of 8
123 strains (hence including *A. vinelandii* DSM 2289 and *B. amyloliquefaciens* LMG 9814 genome
124 sequences) by searching for secondary metabolites gene clusters and Type 3 Secretion Systems.
125 Results are reported in Table 3. As expected, many NonRibosomal Peptide Synthetases (NRPS) gene
126 clusters and genes involved in the biosynthesis of relevant secondary metabolites were found, in
127 particular for strains of genera *Bacillus*, *Pseudomonas*, and *Burkholderia.* Most notably, NRPS were
128 found in all strains. Strains of *Bacillus* were those harbouring the largest number of NRPS.
129 Concerning PGP abilities, siderophore production gene clusters were found for *P.s granadensis*
130 *A23/T3c*, *P. fluorescens DR54*, and *R. acquatilis BB23/T4d*. Finally, several strains contained gene
131 clusters for protection against oxidative damage (as for aryl polyene, carotenoid production) [28],
132 which could contribute to strain survival in soil and in bioinoculant formulation.

133

134 **Table 3.** Presence of gene clusters for production of secondary metabolites and presence of Type 3 Secretion Systems
135 (T3SS). NRPs, nonribosomal peptide synthetases. The number of T3SS is reported.

Strains	NRPSs	Other secondary metabolite predictions	T3SS

***Azotobacter vinelandii* DSM 2289**		arylpolyene	1
***Bacillus amyloliquefaciens* LMG 9814**	macrolactin, fengycin, piliplastin, difficidin, surfactin, bacylisin, baciullibactin	terpene	0
***Bacillus* sp. BV84**	macrolactin, fengycin, piliplastin, difficidin, surfactin, bacylisin, baciullibactin, macrolactin	terpene	0
***Bacillus licheniformis* PS141**	butirosin, fengycin, lichensyn, lassopeptide, surfactin, bacylisin, baciullibactin	terpene	0
***Burkholderia ambifaria* MCI7**	T1PKS, NRPS	terpene, arylpolyene, phosphonate, pyrrolnitrin	0
***Pseudomonas granadensis* A23/T3c**	NRPS, bacillomycin, pyoverdin, fragin	siderophore, terpene, arylpolyene	0
***Pseudomonas fluorescens* DR54**	RiPP-like, pyoverdin, coelibactin, anikasin	siderophore, terpene, arylpolyene	0
***Rahnella aquatilis* BB23/T4d**	lankacidin	betalactone, terpene, arylpolyene	0

Pangenome analysis of the 8 strains was performed by selecting four panels of strains with various degree of phylogenetically relatedness (Supplementary Table S2). This analysis, after sorting out from unique genes for each strain (Supplementary Dataset S1), can allow selected genes candidate for primer design to allow traceability of bioinoculants in the environment.

Discussion

In this study, the whole-genome sequencing and annotation were carried out to provide useful insights into the mechanisms involved in plant growth of six bacterial strains belonging to well-known plant-growth promoting bacterial species as *Bacillus velezensis*, *Bacillus licheniformis*, *Burkholderia ambifaria*, *Pseudomonas granadensis*, *Pseudomonas fluorescens* and *Rahnella aquatilis*, previously selected to set up microbial consortia for their PGP traits . (Tabacchioni et al., 2021). However, the mechanisms of these strains for promoting plant growth and biocontrol are unclear.

151 The genomes of the strains *Azotobacter vinelandii* DSM2289 and *Bacillus amyloliquefaciens* LMG
152 9814, composing the microbial consortia, were previously sequenced, thus, sequencing data were
153 used for genome comparative analysis.
154 The genome size of all sequenced strains was in agreement with previously reported genomes of their
155 related strains in NCBI database. Many gene clusters related to secondary metabolites involved in
156 indirect mechanisms of plant growth have been detected in all sequenced genomes suggesting a high
157 potential of these strains as biocontrol agents. It is noteworthy that the presence of gene clusters
158 putatively involved in pathogenic traits (Type III secretion system) was detected only in *Azotobacter*
159 *vinelandii* DSM2289.
160 Bacteria belonging to *B. licheniformis* species are known to improve the growth several plants among
161 which chrysanthemum, tomato and maize through several mechanisms among which production of
162 Indole Acetic Acid (IAA), siderophore, ammonia, phosphate and zinc solubilisation other than the
163 improvement of water use efficiency (Aktar et al., 2020;de O Nunes et al. 2023; Nilina et al, 2023).
164 *B. licheniformis* PS141, for its plant growth promoting properties (unpublished results) is also a
165 component of the commercial products Bio-Semina PW (Agriges). The sequencing of this strain
166 revealed the presence of several genes involved in the biosynthesis of secondary metabolites involved
167 in indirect mechanisms of plant growth.
168 *B. velezensis* BV84 is an endophytic bacterium isolated from grape leafs with known for its biocontrol
169 and plant growth promoting properties (unpublished results).It is also a component of the commercial
170 bioferilizer MICOSAT F (CCSAosta). Bacteria belonging the *Bacillus velezensis* species are known
171 to promote plant growth and control pathogenic microorganisms by producing some secondary
172 metabolites or antibiotics, and efficient colonization of plants (Adeniji and Babalola, 2019). Auxin
173 production and improved nutrient uptake by bacteria belonging to this species were observed as well
174 as phosphate solubilizing ability Talboys et al., 2014, Kim et al., 2017, Hwangbo et al., 2016). WGS
175 sequencing of *B. velezensis* BV84 highlighted the presences of several genes involved in the

176 biosynthesis of antibacterial and antifungal compounds that could be the main mechanisms involved
177 in the PGP activity of this strain.
178 *Burkholderia ambifaria* MCI 7, isolated from the maize rhizosphere, was found to increase maize
179 plants growth in both uninfested soil and soil infested with *F. moniliforme* (Bevivino et al., 2000,
180 Ciccillo et al., 2002). The presence of the gene clusters encoding for pyrrolnitrin biosynthesis suggest
181 the main involvement of this antifungal compound in the biocontrol activity exerted by this strain.
182 The presence in the genome of this strain of genes related to P utilization (phosphonate) could also
183 suggest a role of this bacterium in the P metabolism. Moreover, the presence of genes encoding for
184 the biosynthesis of compounds (arylpolyene) which could protect from oxidative stress (Schoner et
185 al., 2015) could contribute to strain survival in soil and in bioinoculant formulation.
186 *Pseudomonas fluorescens* DR54, isolated from sugar beet rhizosphere (Nielsen et al., 1998), has
187 proven to be an effective biocontrol agent against the phytopathogenic fungi *Phytium ultimun* and
188 *Rhizoctonia solani* in microcosm, greenhouse and field conditions (Thrane et al., 2000, Thrane et al.,
189 2001, Johansen et al., 2005). It was observed that the antifungal lypolypetide viscosinamide produced
190 by this bacterium (Nielsen et al., 1999), affect the *P. ultimum* zoospore formation in microcosm
191 experiments (Thrane et al., 2000). However, also siderophores and pyrrolnitrin, which gene clusters
192 have been detected by whole genome sequencing could be implicated in the biocontrol activity by
193 this strain.
194 *P. granadensis* A23/T3c and *R. aquatilis* BB23/T4d have been reported to improve the growth of
195 *Sorgum bicolor* in greenhouse conditions (Chiarini et al., 1998). Bacteria belonging to *P. granadensis*
196 species have found to possess several PGP traits such as NH_3, HCN and siderophore production, 1-
197 aminocyclopropane-1-carboxylate deaminase activity (ACC) and antifungal compounds (Barbaccia
198 et al., 2022; Reira et al., 2017). The species *R. aquatilis* has been described to protect plants from a
199 wide range of phytopathogenic organisms in different ways (Kong et al., 2020; Palmieri et al., 2020,
200 Yuan et al., 2020). Genomic analysis of *P. granadensis* A23/Tc and *R. aquatilis* BB23/T4d revealed
201 the great potential of these strains as biocontrol agents as the presence of several gene clusters

involved in biocontrol activity have been detected in their genomes. Moreover, also for these strains the presence of genes involved in the biosynthesis of antioxidans compounds could be important to increase the fitness of these strains as inoculants.

The data reported in this work allow to rationally design SynComs formulation which take into account relevant genomic information, as the putative biosynthetic pathways for relevant secondary metabolites and highlight the presence of a large unique gene pool for each strains, which could be further used to develop molecular methods for biomonitoring (i.e. qPCR). The absence of pathogenicity traits as well as the identification of genes that are likely responsible for the plant-growth promoting features that were experimentally described suggest that SIMBA microbial consortia represent good candidates for the scaling-up process and application as "plant probiotics" for a sustainable agriculture.

Funding

This research was funded by the European Union's Horizon 2020 Research and Innovation Programme under Grant agreement No. 818431 (SIMBA, Sustainable Innovation of Microbiome Applications in Food System).The funding sources had no involvement in study design, collection, analysis and interpretation of data, writing, and decision to submit the paper. This output reflects only the authors' views, and the Research Executive Agency (REA) of the EU cannot be held responsible for any use that may be made of the information contained therein. This project has also received funding from SUS-MIRRI.IT project funded by the European Union─NextGeneration EU, PNRR─Mission 4 "Education and Research" Component 2: from research to business, Investment 3.1: Fund for the realization of an integrated system of research and innovation infrastructures─IR0000005 (D.M. Prot. n.120 del 21/06/2022).

Acknowledgments

The authors are also grateful to the SIMBA partners https://simbaproject.eu/about/consortium/ and the Coordinator of SIMBA Project Dr. Anne Pihlanto (Natural Resources Institute Finland (Luke) for helpful discussion.

Author Contributions: Conceptualization, A.B.; investigation, all authors; writing—original draft preparation, L.C. A.M., A.B.; writing—review and editing, A.B. and A.M. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Oleńska, E.; Małek, W.; Wójcik, M.; Swiecicka, I.; Thijs, S.; Vangronsveld, J. Beneficial features of plant growth-promoting rhizobacteria for improving plant growth and health in challenging conditions: A methodical review. *Sci. Total Environ.* **2020**, *743*, 140682, doi:https://doi.org/10.1016/j.scitotenv.2020.140682.
2. Li, J.; Wang, J.; Liu, H.; Macdonald, C.A.; Singh, B.K. Application of microbial inoculants significantly enhances crop productivity: A meta-analysis of studies from 2010 to 2020. *J. Sustain. Agric. Environ.* **2022**, *1*, 216–225.
3. Mapelli, F.; Mengoni, A.; Riva, V.; Borin, S. Bacterial culturing is crucial to boost sustainable agriculture. *Trends Microbiol.* **2022**, *xx*, 1–4, doi:10.1016/j.tim.2022.10.005.
4. Vaccaro, F.; Cangioli, L.; Mengoni, A.; Fagorzi, C. Synthetic plant microbiota challenges in nonmodel species. *Trends Microbiol.* **2022**, *30*, 922–924.

5. Mącik, M.; Gryta, A.; Frąc, M. Biofertilizers in agriculture: An overview on concepts, strategies and effects on soil microorganisms. *Adv. Agron.* **2020**, *162*, 31–87, doi:10.1016/bs.agron.2020.02.001.
6. Maitra, S.; Brestic, M.; Bhadra, P.; Shankar, T.; Praharaj, S.; Palai, J.B.; Shah, M.M.R.; Barek, V.; Ondrisik, P.; Skalický, M.; et al. Bioinoculants—natural biological resources for sustainable plant production. *Microorganisms* **2022**, *10*, 51, doi:10.3390/microorganisms10010051.
7. Wang, C.; Li, Y.; Li, M.; Zhang, K.; Ma, W.; Zheng, L.; Xu, H.; Cui, B.; Liu, R.; Yang, Y.; et al. Functional assembly of root-associated microbial consortia improves nutrient efficiency and yield in soybean. *J. Integr. Plant Biol.* **2021**, *63*, 1021–1035, doi:10.1111/jipb.13073.
8. Tabacchioni, S.; Passato, S.; Ambrosino, P.; Huang, L.; Caldara, M.; Cantale, C.; Hett, J.; Del Fiore, A.; Fiore, A.; Schlüter, A.; et al. Identification of beneficial microbial consortia and bioactive compounds with potential as plant biostimulants for a sustainable agriculture. *Microorganisms* **2021**, *9*, 1–23, doi:10.3390/microorganisms9020426.
9. Jain, A.; Singh, S.; Kumar Sarma, B.; Bahadur Singh, H. Microbial consortium-mediated reprogramming of defence network in pea to enhance tolerance against Sclerotinia sclerotiorum. *J. Appl. Microbiol.* **2012**, *112*, 537–550, doi:10.1111/j.1365-2672.2011.05220.x.
10. Ke, J.; Wang, B.; Yoshikuni, Y. Microbiome engineering: synthetic biology of plant-associated microbiomes in sustainable agriculture. *Trends Biotechnol.* **2021**, *39*, 244–261.
11. Fagorzi, C.; Passeri, I.; Cangioli, L.; Vaccaro, F.; Mengoni, A. When biodiversity preservation meets biotechnology: The challenge of developing synthetic microbiota for resilient sustainable crop production. *J. Sustain. Agric. Environ.* **2023**, *2*, 5–15, doi:10.1002/sae2.12038.
12. Shayanthan, A.; Ordoñez, P.A.C.; Oresnik, I.J. The Role of Synthetic Microbial Communities (SynCom) in Sustainable Agriculture. *Front. Agron.* **2022**, *4*, 1–13,

doi:10.3389/fagro.2022.896307.

13. Delgado-Baquerizo, M. Simplifying the complexity of the soil microbiome to guide the development of next-generation SynComs. *J. Sustain. Agric. Environ.* **2022**, *1*, 9–15.

14. Vorholt, J.A.; Vogel, C.; Carlström, C.I.; Müller, D.B. Establishing Causality: Opportunities of Synthetic Communities for Plant Microbiome Research. *Cell Host Microbe* **2017**, *22*, 142–155, doi:10.1016/j.chom.2017.07.004.

15. Sessitsch, A.; Pfaffenbichler, N.; Mitter, B. Microbiome Applications from Lab to Field: Facing Complexity. *Trends Plant Sci.* **2019**, *24*, 194–198, doi:10.1016/j.tplants.2018.12.004.

16. Qi, M.; Berry, J.C.; Veley, K.M.; O'Connor, L.; Finkel, O.M.; Salas-González, I.; Kuhs, M.; Jupe, J.; Holcomb, E.; Glavina del Rio, T. Identification of beneficial and detrimental bacteria impacting sorghum responses to drought using multi-scale and multi-system microbiome comparisons. *ISME J.* **2022**, *16*, 1957–1969.

17. Liu, Y.-X.; Qin, Y.; Bai, Y. Reductionist synthetic community approaches in root microbiome research. *Curr. Opin. Microbiol.* **2019**, *49*, 97–102, doi:https://doi.org/10.1016/j.mib.2019.10.010.

18. Hett, J.; Döring, T.F.; Bevivino, A.; Neuhoff, D. Impact of microbial consortia on organic maize in a temperate climate varies with environment but not with fertilization. *Eur. J. Agron.* **2023**, *144*, 126743.

19. Hett, J.; Neuhoff, D.; Döring, T.F.; Masoero, G.; Ercole, E.; Bevivino, A. Effects of multi-species microbial inoculants on early wheat growth and litterbag microbial activity. *Agronomy* **2022**, *12*, 899.

20. Graziano, S.; Caldara, M.; Gullì, M.; Bevivino, A.; Maestri, E.; Marmiroli, N. A Metagenomic and Gene Expression Analysis in Wheat (T. durum) and Maize (Z. mays) Biofertilized with PGPM and Biochar. *Int. J. Mol. Sci.* **2022**, *23*, 10376.

21. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.O.N.; Prjibelski, A.D.; et al. SPAdes: A New Genome

Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477, doi:10.1089/cmb.2012.0021.

22. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069, doi:10.1093/bioinformatics/btu153.

23. Meier-Kolthoff, J.P.; Göker, M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* **2019**, *10*, 1–10.

24. Meier-Kolthoff, J.P.; Auch, A.F.; Klenk, H.P.; Göker, M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **2013**, *14*, doi:10.1186/1471-2105-14-60.

25. Blin, K.; Shaw, S.; Kautsar, S.A.; Medema, M.H.; Weber, T. The antiSMASH database version 3: increased taxonomic coverage and new query features for modular enzymes. *Nucleic Acids Res.* **2021**, *49*, D639–D643.

26. Hui, X.; Chen, Z.; Lin, M.; Zhang, J.; Hu, Y.; Zeng, Y.; Cheng, X.; Ou-Yang, L.; Sun, M.; White, A.P. T3SEpp: an integrated prediction pipeline for bacterial type III secreted effectors. *Msystems* **2020**, *5*, 10–1128.

27. Page, A.J.; Cummins, C.A.; Hunt, M.; Wong, V.K.; Reuter, S.; Holden, M.T.G.; Fookes, M.; Falush, D.; Keane, J.A.; Parkhill, J. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **2015**, *31*, 3691–3693, doi:10.1093/bioinformatics/btv421.

28. Schöner, T.A.; Gassel, S.; Osawa, A.; Tobias, N.J.; Okuno, Y.; Sakakibara, Y.; Shindo, K.; Sandmann, G.; Bode, H.B. Aryl polyenes, a highly abundant class of bacterial natural products, are functionally related to antioxidative carotenoids. *ChemBioChem* **2016**, *17*, 247–253.