Cell Host & Microbe
Review

Elucidating Bacterial Gene Functions in the Plant Microbiome

Asaf Levy,[1,7,*] Jonathan M. Conway,[2,3] Jeffery L. Dangl,[2,3,4,5] and Tanja Woyke[1,6]
[1]DOE Joint Genome Institute, Walnut Creek, CA, USA
[2]Department of Biology, University of North Carolina, Chapel Hill, NC, USA
[3]Howard Hughes Medical Institute, Chevy Chase, MD, USA
[4]The Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC, USA
[5]Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, NC, USA
[6]School of Natural Sciences, University of California, Merced, Merced, CA, USA
[7]Present address: Department of Plant Pathology and Microbiology, The Robert H. Smith Faculty of Agriculture, Food, and Environment, The Hebrew University of Jerusalem, Rehovot 76100, Israel
*Correspondence: alevy@mail.huji.ac.il
https://doi.org/10.1016/j.chom.2018.09.005

There is a growing appreciation for the important roles microorganisms play in association with plants. Microorganisms are drawn to distinct plant surfaces by the nutrient-rich microenvironment, and in turn some of these colonizing microbes provide mutualistic benefits to their host. The development of plant probiotics to increase crop yield and provide plant resistance against biotic and abiotic stresses, while minimizing chemical inputs, would benefit from a deeper mechanistic understanding of plant-microbe interaction. Technological advances in molecular biology and high-throughput -omics provide stepping stones to the elucidation of critical microbiome gene functions that aid in improving plant performance. Here, we review -omics-based approaches that are propelling forward the current understanding of plant-associated bacterial gene functions, and describe how these technologies have helped unravel key bacterial genes and pathways that mediate pathogenic, beneficial, and commensal host interactions.

Introduction

All land plants host a microbiome composed of bacteria, fungi, oomycetes, viruses, archaea, and protists. These organisms inhabit primarily the root environment (rhizosphere; the area immediately adjacent to the root), the rhizoplane (the root surface), and to a lesser extent the leaf (phyllosphere), seed (spermosphere), and internal (endosphere) plant environments. Microbes are attracted to the rich nutrients provided by the plant and are sorted from the surrounding environment (soil, water, and air), presumably by the plant immune system, the exudates that the plant secretes to the soil in the immediate vicinity of the root, and the ability to outcompete other microbes. Interestingly, very different plants, such as the dicot *Arabidopsis* and the monocot barley, share a core root and rhizosphere microbiome composed of mainly Proteobacteria, Actinobacteria, and Bacteroidetes, and this core microbiome is distinct from the bulk soil. Plant microbiome research has rapidly expanded along with the understanding that the microbiome can have far-reaching implications on the plant's health, development, and productivity (Mayak et al., 2004; Niu et al., 2017).

Nonetheless, most plant microbiome studies either use amplicon-based microbial ecology approaches to describe the community structure or focus on a limited set of model plant pathogens or beneficial microorganisms. To improve the mechanistic understanding of the interaction between plants and their microbiomes there needs to be a shift from community structure description to systematic microbial function elucidation. Common microbial ecology tools (e.g., 16S ribosomal RNA or *gyrB* gene amplicon sequencing) provide insights into the makeup of a bacterial community. However, these techniques cannot determine if a certain microbe is harmful, neutral, or beneficial to the plant. These outcomes are dependent on the genetics of both the host and the microbiome. The presence or absence of even a small number of accessory genes in either the plant (e.g., disease resistance, or *R* genes) or its microbiome (e.g., virulence factors, or genes that dampen plant stress responses, modulate plant hormone levels, or mobilize nutrients) may cause a drastic change in the nature of their interaction. Moreover, samples that are different in their species diversity can still encode similar gene functions and proteomes, as shown in four tree species (Lambais et al., 2017) as well as in different samples of the human microbiomes (Lozupone et al., 2012). Research using model microorganisms, such as different root-nodulating rhizobiales and the phytopathogen *Pseudomonas syringae*, has identified factors contributing to mutualism or virulence, respectively (Glick, 2014; Xin et al., 2018). However, mechanistic studies on model plant-associated bacterial isolates tend to ignore the effect of the extant plant microbiome during colonization and persistence of the studied strain. Methodological advances in molecular biology in multiple -omics fields, including genomics, transcriptomics, proteomics, and metabolomics, have begun to yield insights into the functions performed at the community-level by plant-associated bacterial genes and pathways.

Here, we review recent developments in the elucidation of bacterial gene functions and characterization of molecular changes at the plant-bacteria interface through the application of -omics techniques. A thorough molecular understanding of

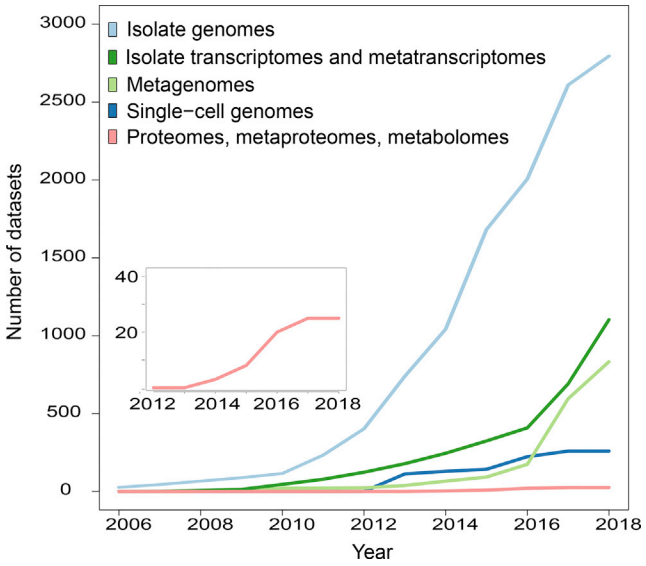

Figure 1. Cumulative Number of Genomes, Metagenomes, Meta(Transcriptomics), Meta(Proteomics), and Metabolomics Datasets from Plant Microbiome Studies per Year; Information Was Obtained from Integrated Microbial Genomes, NCBI BioProject, ProteomeExchange, and MassIVE systems; Inset Shows the Total Number of Proteomes, Metaproteomes, and Metabolomes

plant microbiome functions will have significant agricultural implications, including the deployment of useful microbes and microbial-derived products to increase crop yields. These might include inoculating crops with a supportive and robust microbial community or engineering plants with beneficial microbial genes to confer higher productivity and resistance against plant diseases, pests, and abiotic stresses. Ultimately, these technologies will contribute to more efficient and sustainable agriculture.

Overview of -Omics Approaches to Understand Plant Microbiome Gene Function

Genomics and Metagenomics

The striking reduction of DNA sequencing costs has led to the creation of large-scale bacterial genome collections. Currently, hundreds of public genomic datasets of plant-associated bacterial isolates, single cells, and metagenomes become available each year (Figure 1). High-quality bacterial isolate genomes can be compared to identify candidate genes and pathways that correlate with a given phenotype of interest, such as association with a specific niche, virulence, or a beneficial phenotypic trait. These genes can then be manipulated to test for the predicted function. In recent years, thousands of bacterial isolate genomes have been sequenced from different plant environments and compared to identify bacterial genes that affect general adaptation to plants (Levy et al., 2018), adaptation to root versus shoot (Bai et al., 2015), nodulation and nitrogen fixation (Seshadri et al., 2015), biocontrol activity (Hossain et al., 2015), and quorum sensing (Schaefer et al., 2013).

An alternative to sequencing bacterial isolates is to sequence a plant microbiome metagenome ("shotgun metagenomics"). The "meta" prefix used here, as with other -omics techniques, indicates that the data represent measurements captured from the entire microbial community and not from a single isolate. In metagenomics, genetic sequence information is captured for the many species across a microbiome that cannot be represented by cultivation. Metagenome sequencing projects revealed genes that are enriched in the endosphere (Sessitsch et al., 2012) and rhizospheres of different plants (Bulgarelli et al., 2015; Ofek-Lalzar et al., 2014), elucidated genes that are correlated with biocontrol activity (Mendes et al., 2018), and even led to the discovery of novel metabolic enzymes (Campos et al., 2016). A major challenge in metagenomics is to assemble the sequencing reads into high-quality metagenome-assembled genomes where all genes in a genome are captured and the assembled contigs are assigned to the correct organisms. This may be particularly challenging for rare organisms. Other hurdles include proper taxonomic assignment of the assembled genomes and differentiation between related strains in samples containing a high degree of strain heterogeneity. In endosphere microbiomes, large amounts of host DNA masking the microbial DNA can further complicate shotgun metagenomic approaches.

Functional metagenomics provides an approach to systematically test the effect of gene gain of function. Here, novel genes discovered in metagenomes can be expressed in a heterologous host microbe or *in vitro*, which enables functional assays to be employed to test for novel activities. However, cloning and heterologous expression of some genes may be intractable, or expression in a heterologous host microbe may yield a different phenotype. This approach is often used to discover novel antibiotic biosynthesis or resistance genes within soil metagenomes, but it has not yet been systematically applied to plant metagenomes. One potential application of functional metagenomics could be the systematic identification of novel plant growth-promoting genes by heterologous expression in a root colonizer.

A complementary approach to metagenomics is single-cell sequencing. Here, prior to sequencing, single cells are first isolated and lysed, and the DNA is amplified through a multiple displacement amplification reaction. Single-cell sequencing allows genome sequencing of bacteria that cannot be cultivated, provides access to the genetic makeup of rare taxa, and overcomes the challenge of assigning a DNA sequence to a certain cell, thereby facilitating linkage of plasmids and viruses to their bacterial host. The main limitation of single-cell sequencing is that the resulting genomes are generally less complete, more fragmented, and more susceptible to contamination as compared with sequencing of clonal cultured bacterial isolates.

Genetically tractable microorganisms can be tested for gene function using systematic gene loss-of-function approaches. One powerful approach is transposon sequencing (TnSeq), in which all genes in a genome are mutated by transposon insertion to test their involvement in a given biological process. A variant of TnSeq is randomly barcoded transposon mutagenesis sequencing (RB-TnSeq), in which TnSeq is coupled with random DNA barcoding of each mutant to identify genes that affect microbial fitness under specific growth conditions (Price et al., 2018). This approach was used to mutate the genomes of 33 bacterial strains, some of which are plant-associated, and provided a remarkable repository that includes the mutant phenotypes of 100,000 bacterial genes (Price et al., 2018). TnSeq-based approaches were recently applied to identify bacterial genes involved in *Arabidopsis* and legume root colonization

(Cole et al., 2017; Salas et al., 2017), in bacterial persistence in tomatoes (de Moraes et al., 2017), and in xylose metabolism (Price et al., 2018).

Transcriptomics and Metatranscriptomics

Transcriptomic analysis of plant-associated bacteria using RNA sequencing (RNA-seq) technology, or gene expression microarray approaches, reveals genes that are differentially expressed under certain conditions. To date, most of the plant-associated bacterial transcriptomic studies have been performed by culturing bacteria separate from the plant host. RNA-seq was used, for example, to detect genes responding to the presence of plant extract (Coutinho et al., 2015). The challenge for the study of bacterial transcriptomes *in planta* is that plant transcripts significantly outnumber bacterial transcripts and most bacterial transcripts are housekeeping ribosomal RNAs. Hence, achieving a sufficient concentration of bacterial mRNA transcripts for sequencing and differential expression analysis is difficult. Several *in planta* bacterial isolate transcriptome studies report simultaneous plant and bacterial gene expression (termed "dual RNA-seq") (Pankievicz et al., 2016; Paungfoo-Lonhienne et al., 2016; Roux et al., 2014). Recently, Nobori et al. (2018) developed two highly correlated approaches to significantly enrich for the transcriptome of *P. syringae* in an *Arabidopsis* leaf infection model. In the first, a new isolation buffer that stabilizes the bacterial RNA was used during leaf grinding. This was followed by filtration and centrifugation to separate bacterial cells from plant cells prior to RNA isolation. The second approach used selective depletion of plant-derived transcripts with customized probes. It remains to be seen if these approaches can be applied to root-dwelling bacteria. RNA-seq technology also enables detection of intricate transcriptome regulation such as gene operons, small noncoding RNA, antisense RNA, and riboswitches (Filiatrault et al., 2010).

In metatranscriptomics, transcripts of the entire community are directly sequenced from environmental samples. This allows insight into the transcriptional state of many microorganisms simultaneously. Metatranscriptomics were used, for example, to identify bacterial genes from the rhizosphere that are differentially expressed during *Arabidopsis* development (Chaparro et al., 2014) and invasion by a fungal pathogen (Chapelle et al., 2016).

Decreasing sequencing costs have enabled the increased use of transcriptomics and metatranscriptomics (Figure 1) to gain insights into the dynamics of bacterial gene expression. Transcriptomic analysis enables the dynamics and regulation of actively transcribed genes to be detected, thereby presenting an advantage over genomic analysis. Metatranscriptomics, however, is limited by the fact that transcripts can rarely be assigned to specific microorganisms without high-quality reference genomes. Alternatives to sequencing-based transcriptomic approaches, such as the hybridization-based NanoString technology, may allow improved bacterial transcript detection in mixed plant microbiome transcript samples. As techniques for the enrichment and detection of bacterial transcripts further improve and become applicable to a broad array of plant-bacteria systems, we expect that transcriptomic approaches will transform our understanding of plant-associated bacterial functions.

Proteomics and Metaproteomics

Proteomics and metaproteomics approaches, mostly based on liquid chromatography-tandem mass spectrometry technology, reveal the diversity of bacterial proteins within an environment in a semi-quantitative manner. These techniques involve sample collection, protein extraction, isolation and fractionation, mass spectroscopy analysis, and comparison with a proteome database. Unlike genomics, and to a lesser extent transcriptomics, proteomics measures the functional protein components produced by a cell rather than identifying the potential to make them. Therefore, proteomics approaches provide a more precise snapshot of the active pathways within a sample. (Meta)proteomics has been used to measure the phyllosphere metaproteome of forest trees (Lambais et al., 2017), to detect proteins differentially secreted by plant growth-promoting bacterial (PGPB) strains in response to root exudates (Kierul et al., 2015), and identify the organisms and proteins responsible for nitrogen fixation and methane oxidation in rice fields (Bao et al., 2014). Proteomics can be limited by low protein quality and concentration, low sensitivity due to host proteins and microbial complexity, and *de novo* protein identification if a (meta)genome reference sequence is lacking. The Vorholt lab pioneered the use of metaproteogenomics, in which proteins present in complex microbial communities are identified based on metagenomes generated from plant microbiomes. The approach doubled the number of proteins that could be identified compared with protein identification using public databases alone (Delmotte et al., 2009; Knief et al., 2012).

Unfortunately, the current application of proteomics to describe plant-associated bacterial communities is limited (Figure 1) due to various factors, including relatively low bacterial protein expression levels in complex plant-associated samples and consequent detection limits, and the need for a comprehensive peptide reference database. We hope to see higher use of proteomics in studies examining plant-bacteria interactions in the future, complementing the large number of genomics and transcriptomics studies.

Metabolomics

Various bacterial genes, such as the *Nodulation* (*Nod*) genes that synthesize the Nod factors as part of root nodulation, directly affect the host plant or microbial metabolism. Using targeted or untargeted metabolomics, changes in specific metabolite levels can be measured in response to a given treatment. Recently, metabolomics was used to demonstrate how the chemical exudation from grass (*Avena barbata*) roots over the course of development affects rhizosphere community assembly and succession by enriching for bacteria with substrate preference for the exuded metabolites, mostly aromatic organic acids (Zhalnina et al., 2018).

There are several challenges associated with metabolome analysis in plant-microbe systems such that they have not been widely adopted (Figure 1). Similar to proteomics, the costs, equipment, and technical expertise necessary to perform metabolite studies make them less accessible than DNA sequencing. Further, the sizes of public metabolite reference databases are limited, and it can be difficult to assign a measured metabolite to a specific organism. Nevertheless, metabolomics offers a powerful tool to detect and quantify small molecules and molecular changes at the plant-bacteria interface. Discovery

Table 1. Advantages and Disadvantages of Different -Omics Approaches Employed in Studying Gene Function of Plant-Associated Bacteria

Method	Strengths	Limitations
Whole-genome sequencing of isolate genomes combined with comparative genomics	• Allows identification of genes and genomic features associated with a certain environment or phenotype. • High confidence in the association between different genes within a genome. • High genome quality in terms of completeness, contiguity, and lack of contamination. • Nearly 3,000 plant-associated bacterial isolate genomes are already publicly available which allow easy and accurate comparisons.	• Many bacteria cannot be cultivated in the lab. • Gene transcription/protein expression is unknown
Metagenome sequencing	• No cultivation efforts are required • Yields genes of bacteria that may not be culturable outside of the plant environment • Allows correlation analysis between genes of different organisms from the same environment	• Complex environments, such as soil, yield a metagenome that is difficult to assemble. • The quality of metagenome-assembled genomes is usually lower than the quality of isolate genomes and often represents multiple related strains. • High sequencing depth is needed to cover a highly complex community and recover organisms present at low abundance. • Difficult to perform *in planta.*
Single-cell sequencing	• No cultivation efforts are required • Yields genes of bacteria that may not be culturable individually outside of the plant environment • Allows a confident assignment of genes to a certain organism	• Biases in sorting and lysis result in incomplete recovery of community members. • Single amplified genomes are often incomplete due to biases in DNA amplification from a single cell.
Whole-genome transposon sequencing (TnSeq)	• Allows identification of all genes within a strain required for or inhibiting a certain process	• Many bacteria cannot be cultured in the lab. • Among culturable bacteria, not all can be transformed and mutagenized successfully with the plasmid carrying the transposon. • Cannot identify genes whose function is provided in *trans* in the population.
Transcriptomics of bacterial isolates and metatranscriptomics	• Allows differential expression analysis between different conditions. • Allows simultaneous detection of microbial and host gene expression. • Allows discovery of novel RNA species (noncoding genes, operons, regulatory elements, etc.) and expressed genes. • Easy to perform *in vitro*, for example, in the presence of plant exudates. • Metatranscriptomics: captures transcriptomes for non-culturable bacteria.	• Difficult to perform *in planta* as bacterial mRNA is found in low concentrations. However, new approaches are facilitating *in planta* gene expression for foliar pathogens. • Metatranscriptomics: transcripts can rarely be assigned to specific organisms. Therefore, it is difficult to determine if a pathway/operon is active in a specific microbe (little information about gene co-expression within an organism). • Metatranscriptomics: an extensive genome/metagenome reference is needed to allow read mapping. • Metatranscriptomics: sensitivity is low due to host transcriptome and high community complexity.

(*Continued on next page*)

Table 1. *Continued*

Method	Strengths	Limitations
Proteomics of bacterial isolates and metaproteomics	• Allows semi-quantification of proteins produced in an environment, including enzymes that provide a snapshot of biocatalytic potential. • RNA and protein copy numbers at the single bacterial cell level are uncorrelated and therefore proteome studies may be more accurate for qualitative protein analysis.	• Technical challenges in most steps of the sample preparation process. • Difficult to compare quantity of different proteins. • Difficult to perform *in planta* due to host contamination.
Metabolomics	• Allows identification of the metabolites produced by bacteria/plants in response to each other. This is important as a large fraction of the interactions between these organisms are mediated by small-molecule metabolites. • Minimally biased assessment of diverse compounds.	• Limited size of public reference databases. • Similarity between primary metabolites of plants and microbes make their source (microbe or plant) difficult to determine. • Metabolites with different biological roles may yield similar signals in mass spectrometry measurements.

of microbial small molecules that significantly boost plant health, growth, and resilience to stress remains a high priority target for achieving sustainable agriculture.

Integration of Multiple -Omics Approaches

To further our understanding of bacterial gene functions, it is critical to combine multiple -omics approaches to overcome the different limitations and biases introduced by each technique (Table 1). For example, by combining whole-genome sequencing of a large isolate collection with metatranscriptome sequencing, one can map a higher number of the transcriptome reads to individual isolate reference genomes. This results in a more sensitive differential expression analysis and improved inference about the role of specific isolates and their genes. Combining transcriptomics and/or proteomics with a relevant (meta)genome reference can provide improved insights into the coordinated expression of genomically co-located gene operons, pathogenicity islands, or biosynthetic gene clusters producing secondary metabolites. Metagenomics, metatranscriptomics, and metaproteomics recently provided an elaborate functional landscape of the polymicrobial-host disease interaction during acute oak decline disease (Broberg et al., 2018). The combination of -omics approaches has also led to the identification of the genes and proteins active in legume-rhizobia symbiosis (Delmotte et al., 2010), organisms and pathways active in carbon cycling in grassland soil (Butterfield et al., 2016), genes responding to changes in plant metabolism due to drought (Xu et al., 2018), and genes that are present and expressed in the root environment of two plant species (Ofek-Lalzar et al., 2014). In the last example, the fact that both metagenomics and metatranscriptomics were performed allowed the authors to conclude that most of the genes enriched in the root zones, compared with the soil, are expressed in the root environment, supporting the hypothesis that these genes are derived from active root-associated microbes. By layering -omics techniques to analyze the same plant microbiome system the gene, transcript, protein, and/or metabolite signals can be integrated to gain a clearer and a more reliable picture of the biological phenomena at work.

Genome Engineering to Define Function

CRISPR-Cas9 technology provides the unprecedented ability to accurately and efficiently engineer organisms of interest, including both eukaryotes and prokaryotes. This may have particular relevance to the studies of plant-associated bacteria, which are typically not model organisms with facile engineering platforms. Recently, a CRISPR system was established in the root-colonizing strain *Pseudomonas putida* KT2440 and was combined with single-stranded DNA recombineering, which allowed various DNA manipulations, from single-nucleotide changes to large chromosomal deletions, including simultaneous introduction of multiple mutations (Aparicio et al., 2018). Nonetheless, the new mutations were not tested for an *in planta* phenotype. This is likely the first of many plant-associated strains that will be manipulated using CRISPR technology to achieve desired traits such as plant growth promotion and disease resistance. The ability to precisely engineer plant-associated bacterial genomes could revolutionize our understanding of bacterial gene function. This is particularly important for the many genes annotated as "hypothetical proteins," and currently lacking any function information.

Functions Performed by Plant-Associated Bacterial Genes and Pathways Identified via -Omics Approaches

Virulence and Modulation of Plant Immunity

Among the most studied bacteria in the plant environment are phytopathogens, where there is special interest in genes and pathways that contribute to virulence and interact with the plant immune system. Virulence is frequently mediated through different secreted toxins and virulence proteins, also known as effectors. These are secreted into the plant cell through various bacterial secretion systems and act to subvert plant immune system mechanisms and to facilitate access to water and host-derived nutrients.

Plants use microbe-triggered immunity (MTI) to detect conserved microbe-associated molecular patterns (MAMPs), such as flagellin- and elongation factor Tu-derived peptides. MAMP-encoding genes share an interesting evolutionary

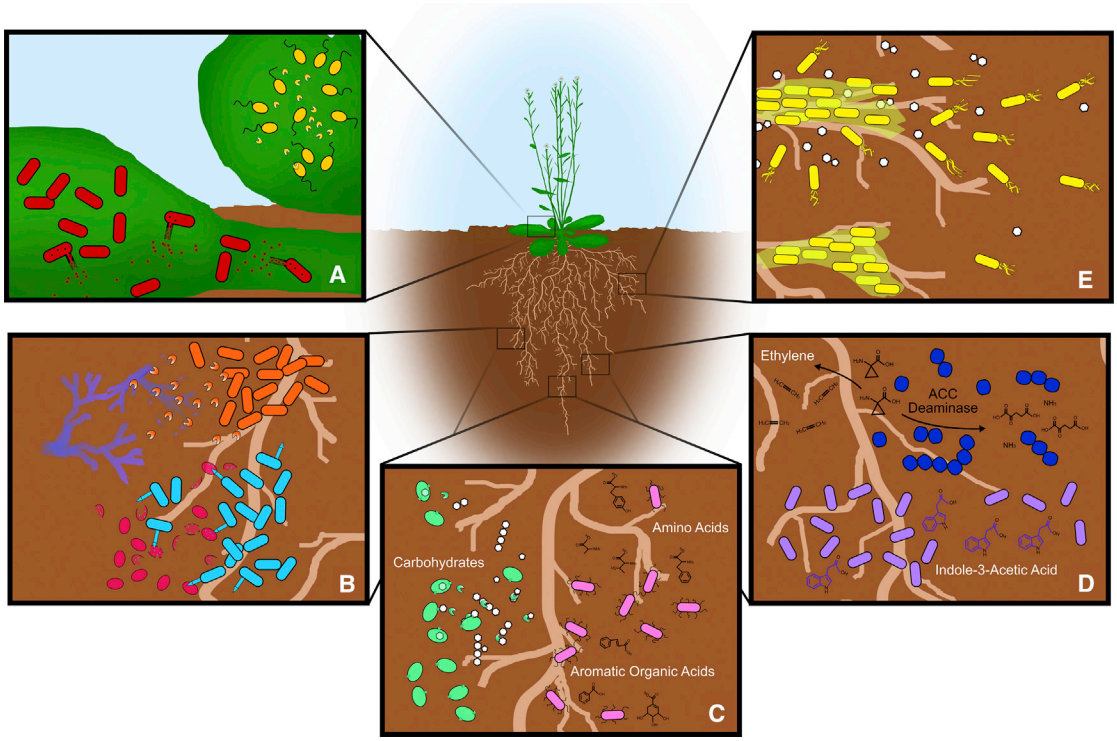

Figure 2. Bacterial Functions Elucidated through -Omics Techniques
(A) Virulence and modulation of plant immunity. Type 3 secretion systems inject type 3 effectors into the plant to affect and evade the immune system. Phytopathogenic bacteria express various factors to gain access to nutrients or respond to plant defenses.
(B) Inter-microbial interactions. Type 6 secretion systems and chitinase production by bacteria mediate bacteria-bacteria and bacteria-fungi antagonism, respectively.
(C) Nutrient uptake. Bacteria consume nutrients exuded by the plant host including carbohydrates, amino acids, and aromatic organic acids.
(D) Symbiosis and plant growth promotion. Bacteria with a 1-aminocyclopropane-1-carboxylate (ACC) deaminase reduce ethylene levels and some bacteria produce plant hormones such as indole-3-acetic acid. Both mechanisms can promote plant growth.
(E) Plant sensing, colonization, and persistence. Colonization of the plant host is driven by bacterial motility, chemotaxis, and biofilm formation.

pattern. They are typically inherent to the lifestyle of the microbe (e.g., flagellin, a significant component of bacterial motility) and are therefore under purifying selection. As MAMP genes need to evade MTI, however, some of their amino acids can be under positive selection (McCann et al., 2012). To suppress MTI, several taxa of Gram-negative bacterial pathogens, including the well-studied model pathogen *P. syringae*, use the needle-like type III secretion system (T3SS) which secretes type III effectors (T3Es) into the plant cell (Xin et al., 2018) (Figure 2A). The T3Es play diverse roles, such as the manipulation of plant transcriptional output to modulate plant immune system response and the establishment of more aqueous apoplast, both of which serve to foster bacterial proliferation (Xin et al., 2018). T3SS and T3Es are transcriptionally upregulated when *P. syringae* is grown *in planta* or on minimal medium compared with growth on a rich medium, and plant MTI suppresses these genes (Nobori et al., 2018). The T3SS in *P. syringae* is under complex regulation that includes, for instance, the two-component system CvsSR, which responds to Ca^{2+} concentrations (Fishman et al., 2017).

Different bacterial transcription factors and quorum sensing systems regulate T3SS and the expression of many downstream virulence factors, including plant cell-wall degrading enzymes, proteases, and proteins defending against plant reactive oxygen species (Broberg et al., 2018; Nobori et al., 2018; Verbon et al., 2017) (Figure 2A). Genes involved in the biosynthesis of phytotoxins, such as syringomycin and syringopeptin, are primarily induced when *P. syringae* grows in the apoplast (Yu et al., 2013). Finally, many bacterial pathogens manipulate different phytohormone biosynthesis pathways to regulate plant growth, development, and defense to provide conditions favorable to their growth and persistence. For instance, metabolomics revealed that the phytopathogen *Rhodococcus fascians* uses the *fas* operon to produce methylated cytokinin to mimic the plant hormone cytokinin (Radhika et al., 2015).

Application of genetic tools that allow a more thorough understanding of bacterial pathogenesis is progressing to define the mechanisms by which root-associated bacteria evade or suppress host MTI responses and to determine how host immune system function contributes to microbiome homeostasis (Garrido-Oter et al., 2018; Hacquard et al., 2017). Notably, the T3SS and its effectors are not limited to phytopathogen genomes, and sequences encoding them are enriched in a metagenome of a healthy barley root and rhizosphere relative to bulk soil (Bulgarelli et al., 2015). It is critical to understand how mechanisms associated with virulence in some taxa are used for symbiotic interactions in others (e.g., the T3SS of different *Rhizobiales*). In addition, comparative genomics of phytopathogens with closely related commensals/mutualists could lead to

the discovery of novel mutualism mechanisms that dampen plant immune response. Complementary approaches to classical genetics, such as systematic mutagenesis or *in planta* bacterial transcriptomics/proteomics, should reveal new bacterial genes involved in successful colonization in the face of a sophisticated immune system.

Inter-microbial Interactions

Plants, and specifically their root environments, are bustling with microorganisms. Consequently, it is not surprising that several of the microbial genes active in the plant environment play a role in cooperative or competitive interactions with other microorganisms (Figure 2B). Some beneficial microbes protect plants against pathogens, serving as biocontrol agents. Direct protection against phytopathogens is accomplished through secretion of antibacterial or antifungal compounds, such as those produced by non-ribosomal peptide synthetases and polyketide synthetases. For example, the polyketide difficidin, encoded by a gene cluster that is conserved among 28 *Bacillus amyloliquefaciens* strains, inhibits the phytopathogen *Xanthomonas axonopodis* pv. *vesicatoria* (Hossain et al., 2015). Recently, a metagenome of a soybean cultivar resistant to the fungal root pathogen *Fusarium oxysporum*, was shown to have higher abundance of predicted antimicrobial biosynthesis genes in the rhizosphere, likely encoding biosynthesis of phenazines and rhamnolipids, compared with a susceptible soybean cultivar or the bulk soil (Mendes et al., 2018). Functional metagenomics of disease-suppressive soil led to the discovery of a novel bacterial chitinase that may confer resistance against fungal phytopathogens (Hjort et al., 2014). In addition, metatranscriptomics revealed that, following exposure to the fungal pathogen *Rhizoctonia solani*, the plant-associated microbiota of sugar beet doubled the expression of stress-related genes, such as those involved in guanosine tetraphosphate metabolism (Chapelle et al., 2016). It is conceivable that plants select and recruit biocontrol strains to complement their defense against pathogens, but a mechanistic understanding of this process is still lacking (Berendsen et al., 2018; Stringlis et al., 2018).

Type VI secretion systems (T6SS) are used by bacteria to introduce toxic antimicrobial proteins into neighboring, mostly bacterial, cells (Figure 2B). T6SS genes were found to be significantly enriched within the barley root and rhizosphere versus bulk soil metagenomes (Bulgarelli et al., 2015). Multiple T6SS-related domains, such as the Hcp effector (pfam05638), are enriched in nearly all plant-associated Proteobacteria reference genomes, including *Alphaproteobacteria*, *Burkholderiales*, and *Xanthomonadaceae* taxa compared with non-plant-associated control genomes (Levy et al., 2018). Moreover, a new family of putative T6SS effectors, named ''Hyde1,'' is specific to *Acidovorax* phytopathogens (absent from *Acidovorax* commensals) and is rapidly evolving within this genus through gene duplication events. A Hyde1-containing isolate was efficient in controlling various leaf bacterial isolates when mixed in culture, whereas a *Hyde1* deletion mutant lost this function, suggesting a role in inter-bacterial competition *in planta* (Levy et al., 2018).

Quorum sensing (QS) is another mechanism by which bacteria interact with one another to coordinate changes in behavior, including activation of virulence, in response to population density. Metagenomic analysis showed that the endophytic rice microbiome encodes at least three types of QS systems: acyl-homoserine lactone (AHL), autoinducer-2 system (AI-2), and the diffusible signal factor system (Sessitsch et al., 2012). AHL QS is active in 40% of 129 poplar rhizosphere and endosphere bacterial isolates tested and in certain strains the QS system lacks some genes and may be used to respond to plant signals (Schaefer et al., 2013). Both T6SS and QS genes were enriched in the rhizoplane communities of wheat and cucumber (Ofek-Lalzar et al., 2014). Other studies have demonstrated that QS molecules can prompt crosstalk between plants and bacteria. For instance, *Medicago truncatula* can detect nanomolar concentrations of AHL from symbiotic and pathogenic bacteria, leading to changes in the plant proteome and the secretion of QS-mimicking compounds (Mathesius et al., 2003).

Microbe-microbe interactions in plant microbiota likely play determinative roles in colonization of the plant host, microbiome succession over plant development, and the ability of mutualists to suppress pathogen growth. The potential to utilize antagonistic microbe-microbe interactions, either in native or engineered strains, as biocontrol agents against plant disease makes research in this area of significant interest. The growing use of small synthetic microbial communities comprised of sequenced isolates, along with precise reductionist -omics approaches applied to these communities will likely yield more predictive models to inter-microbial behavior within and surrounding the plant host.

Uptake of Plant Metabolites

Plants serve as rich sources of nutrients to the bacteria and other microbes living within their tissues and around them (Figure 2C). Through the process of rhizodeposition, plants release organic compounds into the soil. These compounds can contribute to the accumulation of 1.4% of plant-derived carbon in the surrounding microbial biomass (Pausch and Kuzyakov, 2018). For instance, flavonoids, which are carbon-rich secondary metabolites, are enriched in both the *Arabidopsis* and maize rhizosphere, compared with control soil without roots (Petriacq et al., 2017). Bacterial adaptation to plant root exudates is clear at all levels, from the genome to the metabolome. One obvious adaptation observed through comparative genomics is the enrichment of carbohydrate metabolism and transport genes, along with their transcriptional regulators in genomes of phylogenetically diverse plant-associated bacteria (Levy et al., 2018). It is unclear whether carbohydrate metabolism and transport genes are enriched in root environments compared with unplanted soil (Bai et al., 2015; Bulgarelli et al., 2015; Levy et al., 2018; Mendes et al., 2014). However, at least nine such genes are required for *Arabidopsis* root colonization by *P. simiae* (Cole et al., 2017). Methylotrophs and methanotrophs are common in phyllosphere and anoxic rhizosphere environments of grasses, respectively, where they consume reduced one-carbon compounds (Bao et al., 2014; Butterfield et al., 2016; Delmotte et al., 2009; Knief et al., 2012). For instance, metaproteomics analysis revealed the high abundance of methane oxidation proteins from the alphaproteobacterial *Methylocystaceae* family in root tissues of field-grown rice (Bao et al., 2014).

The genomes of many root-associated microbes encode enzymes for degrading plant-derived carbohydrates including cellulose, pectin, xylan, mannan, glucan, and arabinan, and some of these enzymes are secreted in response to root exudates (Campos et al., 2016; Kierul et al., 2015; Ofek-Lalzar et al., 2014;

Sessitsch et al., 2012) (Figure 2C). Notably, certain proteins that are predicted to be secreted to metabolize different carbohydrates are shared between plant-associated bacteria and fungi, despite their large evolutionary distance (Levy et al., 2018). These plant-derived carbons that are imported into the bacterial cell need to be stored. RNA-seq experiments found that the biosynthesis pathway that generates poly-β-hydroxybutyrate for carbon storage is transcriptionally upregulated in the microbiota of wheat and sugarcane root environments (Pankievicz et al., 2016; Paungfoo-Lonhienne et al., 2016).

Plant-associated microbial community structure and metabolism are altered during development and environmental changes, as the nutrients provided by the plants change in composition. Metabolomics and metatranscriptomics experiments revealed that during drought the cereal crop *Sorghum bicolor* accumulates glycerol-3-phosphate (G3P) within its roots, which likely leads to upregulation of G3P ABC-type transporters in the rhizosphere (Xu et al., 2018). Bacteria that accumulate during the development of the wild oat *Avena barbata* prefer to uptake aromatic organic acids exuded by plants such as nicotinic, shikimic, and cinnamic acids and the plant hormones salicylic acid and indole-3-acetic acid (Zhalnina et al., 2018) (Figure 2C).

Interestingly, TnSeq data identified that mutations in nearly 50 bacterial genes required for amino acid metabolism confer a fitness advantage in root colonization versus wild-type bacteria, suggesting that auxotrophy for certain amino acids that are exuded from roots may provide advantages during colonization (Cole et al., 2017). However, genes for the biosynthesis of some of these amino acids, including methionine and branch-chain amino acids, were required for colonization of tomato by *Salmonella* (de Moraes et al., 2017) and alfalfa root by *Ensifer* (Salas et al., 2017). Thus, more experimental data are needed to understand the utility of auxotrophy of different bacterial taxa during plant colonization. A microarray study revealed that *P. syringae* phenylalanine catabolism genes are induced in *Arabidopsis* leaves, which may serve as a means for the bacteria to inhibit plant production of phenylalanine-based defense compounds (Yu et al., 2013).

All bacteria, regardless of their impact on plant health, share a requirement for iron. Iron scavenging using siderophores are induced as part of virulence, and in turn bacterial iron scavenging is modulated by the plant immune system to inhibit microbial proliferation (Broberg et al., 2018; Nobori et al., 2018; Verbon et al., 2017). Iron acquisition and metabolism genes are enriched to some extent in soybean rhizosphere compared with bulk soil, even five years after cultivation (Mendes et al., 2014). Further, metagenomics of rice endophyte communities revealed a high number of genes responsible for siderophore biosynthesis and proteins used for uptake and storage of iron (Sessitsch et al., 2012).

To conclude, elucidation of the nutrients exuded by different plants, the specific bacterial pathways responding to them, and the mechanisms by which bacteria activate or repress further nutrient release, is critical to understanding the bacterial root colonization process.

Symbiosis and Plant Growth Promotion

PGPB are of high interest in agriculture as a means to confer resistance to biotic and abiotic stresses to crops and to increase plant biomass. The molecular mechanisms of root nodulation and biological nitrogen fixation are well studied. *Nod* and *Nif* genes responsible for root nodulation and biological nitrogen fixation, respectively, are enriched in the genomes of plant-associated bacteria and root nodule bacteria in particular (Levy et al., 2018; Seshadri et al., 2015). In wheat and rice roots there is abundant gene expression of certain *Nif* genes with corresponding protein production (Bao et al., 2014; Pankievicz et al., 2016; Sessitsch et al., 2012).

By coupling laser-capture microdissection with RNA-seq, Roux et al. (2014) were able to assess *Medicago* plant and *Sinorhizobium meliloti* gene expression in five root nodule regions. Genes related to cell division, DNA replication, and cell-cycle control are downregulated in nodule areas where bacteria are transitioning toward bacteroids, with flagellar genes being surprisingly active in areas of differentiated and likely non-motile bacteroids. Despite the reported suppression of gene activity, the transcriptome and proteome of *Bradyrhizobium japonicum* during symbiosis with its soybean host shows that nearly 50% of the bacterial genes are expressed in the bacteroid state, including most of the genes for carbon and nitrogen metabolism (Delmotte et al., 2010). Biological nitrogen fixation is performed under anoxic or microaerobic conditions. Indeed, the transcriptome of the nitrogen-fixing *Burkholderia* Q208 strain grown on sugarcane roots provides evidence for energy production in low oxygen conditions through an over 100-fold upregulation of genes active in two anaerobic processes compared with growth in the absence of the plant (Paungfoo-Lonhienne et al., 2016).

Some bacteria are considered PGPB as they combine efficient plant colonization with other attributes beneficial to the plant, including nutrient provision (iron, nitrogen, or phosphorous), release of plant growth hormones (such as indole-3-acetic acid), modulation of ethylene by the enzyme 1-aminocyclopropane-1-carboxylate deaminase (Glick, 2014; Sessitsch et al., 2012), or protection against phytopathogens (Figure 2D). Recently, an operon found in multiple symbiotic rhizobia was shown to biosynthesize the plant hormone gibberellin (Nett et al., 2017). The PGPB *Bacillus amyloliquefaciens* primes plant innate immunity, thereby inducing systemic resistance to secondary infection. This bacterium increases secretion of acetolactate synthase in the presence of root exudates (Kierul et al., 2015), which is involved in systemic resistance through production of the volatile molecule acetoin. Notably, despite the high interest in PGPB, only a handful of known bacterial molecular pathways underpin reproducible plant growth promotion and disease resistance effects. For example, recently, a large group of Rhizobiales were shown to promote *Arabidopsis* root growth, and a few of these strains can also independently override root growth inhibition co-induced with MTI, but the genes underlying these phenotypes are yet unknown (Garrido-Oter et al., 2018). Large-scale comparative -omics between bacteria with and without PGP effects or TnSeq experiments may facilitate identification of novel genetic mechanisms in the future.

Plant Sensing, Colonization, and Persistence

All soil-dwelling microbes that associate with plants, independent of their lifestyle (pathogen, commensal, or beneficial), need to detect the presence of plants. Chemotaxis is used to sense and react to compounds found within the root exudates, such as organic acids and sugars. A chemotaxis response

involves a signal molecule, a chemoreceptor (such as the methyl-accepting chemotaxis protein [MCP]), a cytoplasmic signal transduction system, and a response regulator that controls flagellar or pili activity. The expression of bacterial chemotaxis genes is upregulated in wheat roots, *Arabidopsis* leaf epiphytes, and during later stages of *Arabidopsis* development (Chaparro et al., 2014; Pankievicz et al., 2016; Yu et al., 2013). The endophytic bacterium *Herbaspirilium seropedicae* upregulates MCP gene expression by nearly 20-fold when cells are attached to wheat roots, as revealed by RNA-seq (Pankievicz et al., 2016). Once a signal is perceived by a plant-associated bacterium, it moves toward the plant, primarily through the use of a flagellum (Figure 2E).

The flagellum is likely one of the most crucial components for plant association and it may explain why the flagellin serves as a MAMP to activate the plant immune system. Different flagellar (flagellin, *Fli*, and *Flg*) and chemotaxis (e.g., MCP signaling domain) genes and protein domains are enriched in genomes of plant-associated bacteria from diverse taxa of Proteobacteria and Actinobacteria in comparison with non-plant-associated bacteria of the same taxonomies (Levy et al., 2018). Cell motility is required for root colonization (Cole et al., 2017; Salas et al., 2017), and, along with chemotaxis, these functions are enriched in the root environment (Bai et al., 2015; Knief et al., 2012; Levy et al., 2018; Mendes et al., 2018; Ofek-Lalzar et al., 2014; Sessitsch et al., 2012), as well as the stem and phyllosphere of different plants (Broberg et al., 2018; Delmotte et al., 2009; Lambais et al., 2017). For example, comparative genomics of 206 *Arabidopsis* root-derived versus 33 soil-derived isolates identified a significantly higher fraction of motility genes in the genomes of the root-derived isolates (Bai et al., 2015). To allow colonization, beneficial rhizobacteria evolved mechanisms to suppress flagellin-dependent MAMP-triggered immunity, but many of these mechanisms have yet to be deciphered (Garrido-Oter et al., 2018; Millet et al., 2010). In the phyllosphere, microarray transcriptomic experiments showed that flagellar gene expression is higher among the epiphytes compared with apoplastic bacteria (Yu et al., 2013). Flagellar-mediated motility is regulated in response to the presence of plant material and during abiotic stress such as drought (Coutinho et al., 2015; Xu et al., 2018). Once reaching the plant, bacteria adhere to plant surfaces by producing exopolysaccharides, upregulating adhesins, forming biofilms, and downregulating the flagellum and possibly shifting to twitching motility mediated by type IV pili (Pankievicz et al., 2016; Paungfoo-Lonhienne et al., 2016) (Figure 2E).

Another critical component of bacterial plant colonization and persistence is evasion of plant antimicrobials, such as phytoalexins. Metabolomics and metagenomics of *Arabidopsis* roots showed how certain microbes are resistant to these secreted antimicrobial compounds, whereas others are susceptible, leading to the assembly of a specific rhizosphere community (Stringlis et al., 2018).

There are many key open questions regarding bacterial colonization and subsequent persistence within plant-associated microbiomes. The combination of plant metabolite surveys (using exometabolomics) and bacterial gene activity assays (using transcriptomics/proteomics) could identify candidate genes and pathways that are responsible for plant sensing prior to colonization. Another "*tierra incognita*" is the set of molecular pathways that endophytes require to persist in the plant environment and avoid clearance by the plant innate immune system. Here, enrichment of endophytic RNA/proteins combined with transcriptomics/proteomics could provide candidate sequences for validation through mutagenesis to detect the genes and pathways responsible for persistence.

Summary

The continued development and application of various -omics approaches are beginning to facilitate genome-wide function elucidation in plant-associated bacteria. These methods are useful for detecting candidate genes, proteins, and molecules participating in biological processes, allowing for hypothesis formulation regarding the effect of these molecular entities on a plant phenotype. Clearly, each -omics approach suffers from limitations that affect its sensitivity or specificity (Table 1). By combining multiple -omics approaches some of these limitations can be overcome. However, no -omics technique provides genetic evidence for causality. It is therefore critical to translate the results from -omics studies into testable models, and to follow up experimentally using genetics, biochemistry, and cell biology combined with re-colonization studies using single isolates or defined bacterial synthetic communities. High-throughput loss-of-function (e.g., TnSeq) and gain-of-function (e.g., functional metagenomics) analysis, can facilitate the causal assignment of genes to functions when a phenotype is easily differentiable. Further, improvements in the genetic manipulation of non-model organisms (e.g., with CRISPR-Cas9 technologies) will enable the establishment of connections between genes and functions. Lastly, development and application of probabilistic approaches to study bacterial gene function can aid in function assignment (Plata et al., 2012).

More robust -omics experiments are needed to formulate reproducible and testable models. These research avenues will benefit from the integration of multiple plants and bacterial strains or synthetic communities, various environments, multiple independent assays, and a higher number of biological replicates. We also believe that more focus should be given to the synthesis of genomics and chemistry to identify genes responsible to produce small molecules involved in plant microbiome functions. Identification of these genes, through reverse or forward genetics, can guide the mass production (via strain engineering) of new antimicrobials or compounds that relieve abiotic stress from crops that can be deployed in the field. Other variables that are typically not considered in -omics studies are spatial distribution within plant tissues and inter-kingdom microbial interactions. Therefore, future studies should consider these aspects to provide a more accurate description of plant-microbe interactions. We are still far from having a complete description of plant microbiome function for even the most studied model plants and bacteria. However, the molecular techniques described here are expected to rapidly increase the pace of understanding plant-microbe interactions at the system level.

ACKNOWLEDGMENTS

The work conducted by the US Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under contract no. DE-AC02-05CH11231. J.L.D. is supported by NSF INSPIRE grant IOS-1343020, and by the Office of Science (BER), US Department of Energy, grant no.

DE-SC0014395. J.L.D. is an Investigator of HHMI. We thank Julia A. Vorholt, Susannah G. Tringe, Sarah Lebeis, Sarah R. Grant, and Sharon L. Doty for critical reading of the manuscript.

REFERENCES

Aparicio, T., de Lorenzo, V., and Martinez-Garcia, E. (2018). CRISPR/Cas9-based counterselection boosts recombineering efficiency in *Pseudomonas putida*. Biotechnol. J. *13*, e1700161.

Bai, Y., Muller, D.B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M., Dombrowski, N., Munch, P.C., Spaepen, S., Remus-Emsermann, M., et al. (2015). Functional overlap of the *Arabidopsis* leaf and root microbiota. Nature *528*, 364–369.

Bao, Z., Okubo, T., Kubota, K., Kasahara, Y., Tsurumaru, H., Anda, M., Ikeda, S., and Minamisawa, K. (2014). Metaproteomic identification of diazotrophic methanotrophs and their localization in root tissues of field-grown rice plants. Appl. Environ. Microbiol. *80*, 5043–5052.

Berendsen, R.L., Vismans, G., Yu, K., Song, Y., de Jonge, R., Burgman, W.P., Burmølle, M., Herschend, J., Bakker, P.A.H.M., and Pieterse, C.M.J. (2018). Disease-induced assemblage of a plant-beneficial bacterial consortium. ISME J. *12*, 1496–1507.

Broberg, M., Doonan, J., Mundt, F., Denman, S., and McDonald, J.E. (2018). Integrated multi-omic analysis of host-microbiota interactions in acute oak decline. Microbiome *6*, 21.

Bulgarelli, D., Garrido-Oter, R., Munch, P.C., Weiman, A., Droge, J., Pan, Y., McHardy, A.C., and Schulze-Lefert, P. (2015). Structure and function of the bacterial root microbiota in wild and domesticated barley. Cell Host Microbe *17*, 392–403.

Butterfield, C.N., Li, Z., Andeer, P.F., Spaulding, S., Thomas, B.C., Singh, A., Hettich, R.L., Suttle, K.B., Probst, A.J., Tringe, S.G., et al. (2016). Proteogenomic analyses indicate bacterial methylotrophy and archaeal heterotrophy are prevalent below the grass root zone. PeerJ. *4*, e2687.

Campos, B.M., Liberato, M.V., Alvarez, T.M., Zanphorlin, L.M., Ematsu, G.C., Barud, H., Polikarpov, I., Ruller, R., Gilbert, H.J., Zeri, A.C.d.M., et al. (2016). A novel carbohydrate-binding module from sugar cane soil metagenome featuring unique structural and carbohydrate affinity properties. J. Biol. Chem. *291*, 23734–23743.

Chaparro, J.M., Badri, D.V., and Vivanco, J.M. (2014). Rhizosphere microbiome assemblage is affected by plant development. ISME J. *8*, 790–803.

Chapelle, E., Mendes, R., Bakker, P.A.H., and Raaijmakers, J.M. (2016). Fungal invasion of the rhizosphere microbiome. ISME J. *10*, 265–268.

Cole, B.J., Feltcher, M.E., Waters, R.J., Wetmore, K.M., Mucyn, T.S., Ryan, E.M., Wang, G., Ul-Hasan, S., McDonald, M., Yoshikuni, Y., et al. (2017). Genome-wide identification of bacterial plant colonization genes. PLoS Biol. *15*, e2002860.

Coutinho, B.G., Licastro, D., Mendonca-Previato, L., Camara, M., and Venturi, V. (2015). Plant-influenced gene expression in the rice endophyte *Burkholderia kururiensis* M130. Mol. Plant Microbe Interact. *28*, 10–21.

de Moraes, M.H., Desai, P., Porwollik, S., Canals, R., Perez, D.R., Chu, W., McClelland, M., and Teplitski, M. (2017). *Salmonella* persistence in tomatoes requires a distinct set of metabolic functions identified by transposon insertion sequencing. Appl. Environ. Microbiol. *83*, e03028.

Delmotte, N., Ahrens, C.H., Knief, C., Qeli, E., Koch, M., Fischer, H.-M., Vorholt, J.A., Hennecke, H., and Pessi, G. (2010). An integrated proteomics and transcriptomics reference data set provides new insights into the *Bradyrhizobium japonicum* bacteroid metabolism in soybean root nodules. Proteomics *10*, 1391–1400.

Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., von Mering, C., and Vorholt, J.A. (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. Proc. Natl. Acad. Sci. USA *106*, 16428–16433.

Filiatrault, M.J., Stodghill, P.V., Bronstein, P.A., Moll, S., Lindeberg, M., Grills, G., Schweitzer, P., Wang, W., Schroth, G.P., Luo, S., et al. (2010). Transcriptome analysis of *Pseudomonas syringae* identifies new genes, noncoding RNAs, and antisense activity. J. Bacteriol. *192*, 2359–2372.

Fishman, M.R., Zhang, J., Bronstein, P.A., Stodghill, P., and Filiatrault, M.J. (2017). The Ca2+induced two-component system, CvsSR regulates the type III secretion system and the extracytoplasmic function sigma-factor AlgU in *Pseudomonas syringae* pv. tomato DC3000. J. Bacteriol. *200*, e00538–17.

Garrido-Oter, R., Nakano, R.T., Dombrowski, N., Ma, K.W., AgBiome, T., McHardy, A.C., and Schulze-Lefert, P. (2018). Modular traits of the rhizobiales root microbiota and their evolutionary relationship with symbiotic rhizobia. Cell Host Microbe *24*, 155–167.e5.

Glick, B.R. (2014). Bacteria with ACC deaminase can promote plant growth and help to feed the world. Microbiol. Res. *169*, 30–39.

Hacquard, S., Spaepen, S., Garrido-Oter, R., and Schulze-Lefert, P. (2017). Interplay between innate immunity and the plant microbiota. Annu. Rev. Phytopathol. *55*, 565–589.

Hjort, K., Presti, I., Elväng, A., Marinelli, F., and Sjöling, S. (2014). Bacterial chitinase with phytopathogen control capacity from suppressive soil revealed by functional metagenomics. Appl. Microbiol. Biotechnol. *98*, 2819–2828.

Hossain, M.J., Ran, C., Liu, K., Ryu, C.-M., Rasmussen-Ivey, C.R., Williams, M.A., Hassan, M.K., Choi, S.-K., Jeong, H., Newman, M., et al. (2015). Deciphering the conserved genetic loci implicated in plant disease control through comparative genomics of *Bacillus amyloliquefaciens* subsp. *plantarum*. Front. Plant Sci. *6*, 631.

Kierul, K., Chen, X.-H., Voigt, B., Carvalhais, L.C., Albrecht, D., and Borriss, R. (2015). Influence of root exudates on the extracellular proteome of the plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. Microbiology *161*, 131–147.

Knief, C., Delmotte, N., Chaffron, S., Stark, M., Innerebner, G., Wassmann, R., von Mering, C., and Vorholt, J.A. (2012). Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. ISME J. *6*, 1378–1390.

Lambais, M.R., Barrera, S.E., Santos, E.C., Crowley, D.E., and Jumpponen, A. (2017). Phyllosphere metaproteomes of trees from the Brazilian Atlantic forest show high levels of functional redundancy. Microb. Ecol. *73*, 123–134.

Levy, A., Salas Gonzalez, I., Mittelviefhaus, M., Clingenpeel, S., Herrera Paredes, S., Miao, J., Wang, K., Devescovi, G., Stillman, K., Monteiro, F., et al. (2018). Genomic features of bacterial adaptation to plants. Nat. Genet. *50*, 138–150.

Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., and Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. Nature *489*, 220–230.

Mathesius, U., Mulders, S., Gao, M., Teplitski, M., Caetano-Anolles, G., Rolfe, B.G., and Bauer, W.D. (2003). Extensive and specific responses of a eukaryote to bacterial quorum-sensing signals. Proc. Natl. Acad. Sci. USA *100*, 1444–1449.

Mayak, S., Tirosh, T., and Glick, B.R. (2004). Plant growth-promoting bacteria confer resistance in tomato plants to salt stress. Plant Physiol. Biochem. *42*, 565–572.

McCann, H.C., Nahal, H., Thakur, S., and Guttman, D.S. (2012). Identification of innate immunity elicitors using molecular signatures of natural selection. Proc. Natl. Acad. Sci. USA *109*, 4215–4220.

Mendes, L.W., Kuramae, E.E., Navarrete, A.A., van Veen, J.A., and Tsai, S.M. (2014). Taxonomical and functional microbial community selection in soybean rhizosphere. ISME J. *8*, 1577–1587.

Mendes, L.W., Raaijmakers, J.M., de Hollander, M., Mendes, R., and Tsai, S.M. (2018). Influence of resistance breeding in common bean on rhizosphere microbiome composition and function. ISME J. *12*, 212–224.

Millet, Y.A., Danna, C.H., Clay, N.K., Songnuan, W., Simon, M.D., Werck-Reichhart, D., and Ausubel, F.M. (2010). Innate immune responses activated in *Arabidopsis* roots by microbe-associated molecular patterns. Plant Cell *22*, 973–990.

Nett, R.S., Montanares, M., Marcassa, A., Lu, X., Nagel, R., Charles, T.C., Hedden, P., Rojas, M.C., and Peters, R.J. (2017). Elucidation of gibberellin biosynthesis in bacteria reveals convergent evolution. Nat. Chem. Biol. *13*, 69–74.

Niu, B., Paulson, J.N., Zheng, X., and Kolter, R. (2017). Simplified and representative bacterial community of maize roots. Proc. Natl. Acad. Sci. USA *114*, E2450–E2459.

Nobori, T., Velásquez, A.C., Wu, J., Kvitko, B.H., Kremer, J.M., Wang, Y., He, S.Y., and Tsuda, K. (2018). Transcriptome landscape of a bacterial pathogen under plant immunity. Proc. Natl. Acad. Sci. USA *115*, E3055–E3064.

Ofek-Lalzar, M., Sela, N., Goldman-Voronov, M., Green, S.J., Hadar, Y., and Minz, D. (2014). Niche and host-associated functional signatures of the root surface microbiome. Nat. Commun. *5*, 4950.

Pankievicz, V.C.S., Camilios-Neto, D., Bonato, P., Balsanelli, E., Tadra-Sfeir, M.Z., Faoro, H., Chubatsu, L.S., Donatti, L., Wajnberg, G., Passetti, F., et al. (2016). RNA-seq transcriptional profiling of *Herbaspirillum seropedicae* colonizing wheat (*Triticum aestivum*) roots. Plant Mol. Biol. *90*, 589–603.

Paungfoo-Lonhienne, C., Lonhienne, T.G.A., Yeoh, Y.K., Donose, B.C., Webb, R.I., Parsons, J., Liao, W., Sagulenko, E., Lakshmanan, P., Hugenholtz, P., et al. (2016). Crosstalk between sugarcane and a plant-growth promoting *Burkholderia* species. Sci. Rep. *6*, 37389.

Pausch, J., and Kuzyakov, Y. (2018). Carbon input by roots into the soil: quantification of rhizodeposition from root to ecosystem scale. Glob. Chang. Biol. *24*, 1–12.

Petriacq, P., Williams, A., Cotton, A., McFarlane, A.E., Rolfe, S.A., and Ton, J. (2017). Metabolite profiling of non-sterile rhizosphere soil. Plant J. *92*, 147–162.

Plata, G., Fuhrer, T., Hsiao, T.L., Sauer, U., and Vitkup, D. (2012). Global probabilistic annotation of metabolic networks enables enzyme discovery. Nat. Chem. Biol. *8*, 848–854.

Price, M.N., Wetmore, K.M., Waters, R.J., Callaghan, M., Ray, J., Liu, H., Kuehl, J.V., Melnyk, R.A., Lamson, J.S., Suh, Y., et al. (2018). Mutant phenotypes for thousands of bacterial genes of unknown function. Nature *557*, 503–509.

Radhika, V., Ueda, N., Tsuboi, Y., Kojima, M., Kikuchi, J., Kudo, T., and Sakakibara, H. (2015). Methylated cytokinins from the phytopathogen *Rhodococcus fascians* mimic plant hormone activity. Plant Physiol. *169*, 1118–1126.

Roux, B., Rodde, N., Jardinaud, M.-F., Timmers, T., Sauviac, L., Cottret, L., Carrère, S., Sallet, E., Courcelle, E., Moreau, S., et al. (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.

Salas, M.E., Lozano, M.J., Lopez, 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. (2017). Specificity traits consistent with legume-rhizobia coevolution displayed by *Ensifer meliloti* rhizosphere colonization. Environ. Microbiol. *19*, 3423–3438.

Schaefer, A.L., Lappala, C.R., Morlen, R.P., Pelletier, D.A., Lu, T.-Y.S., Lankford, P.K., Harwood, C.S., and Greenberg, E.P. (2013). LuxR- and luxI-type quorum-sensing circuits are prevalent in members of the *Populus deltoides* microbiome. Appl. Environ. Microbiol. *79*, 5745–5752.

Seshadri, R., Reeve, W.G., Ardley, J.K., Tennessen, K., Woyke, T., Kyrpides, N.C., and Ivanova, N.N. (2015). Discovery of novel plant interaction determinants from the genomes of 163 root nodule bacteria. Sci. Rep. *5*, 16825.

Sessitsch, A., Hardoim, P., Döring, J., Weilharter, A., Krause, A., Woyke, T., Mitter, B., Hauberg-Lotte, L., Friedrich, F., Rahalkar, M., et al. (2012). Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. Mol. Plant Microbe Interact. *25*, 28–36.

Stringlis, I.A., Yu, K., Feussner, K., de Jonge, R., Van Bentum, S., Van Verk, M.C., Berendsen, R.L., Bakker, P.A.H.M., Feussner, I., and Pieterse, C.M.J. (2018). MYB72-dependent coumarin exudation shapes root microbiome assembly to promote plant health. Proc. Natl. Acad. Sci. USA *115*, E5213–E5222.

Verbon, E.H., Trapet, P.L., Stringlis, I.A., Kruijs, S., Bakker, P.A.H.M., and Pieterse, C.M.J. (2017). Iron and immunity. Annu. Rev. Phytopathol. *55*, 355–375.

Xin, X.-F., Kvitko, B., and He, S.Y. (2018). *Pseudomonas syringae*: what it takes to be a pathogen. Nat. Rev. Microbiol. *16*, 316–328.

Xu, L., Naylor, D., Dong, Z., Simmons, T., Pierroz, G., Hixson, K.K., Kim, Y.-M., Zink, E.M., Engbrecht, K.M., Wang, Y., et al. (2018). Drought delays development of the sorghum root microbiome and enriches for monoderm bacteria. Proc. Natl. Acad. Sci. USA *115*, E4284–E4293.

Yu, X., Lund, S.P., Scott, R.A., Greenwald, J.W., Records, A.H., Nettleton, D., Lindow, S.E., Gross, D.C., and Beattie, G.A. (2013). Transcriptional responses of *Pseudomonas syringae* to growth in epiphytic versus apoplastic leaf sites. Proc. Natl. Acad. Sci. USA *110*, E425–E434.

Zhalnina, K., Louie, K.B., Hao, Z., Mansoori, N., da Rocha, U.N., Shi, S., Cho, H., Karaoz, U., Loqué, D., Bowen, B.P., et al. (2018). Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. Nat. Microbiol. *3*, 470–480.