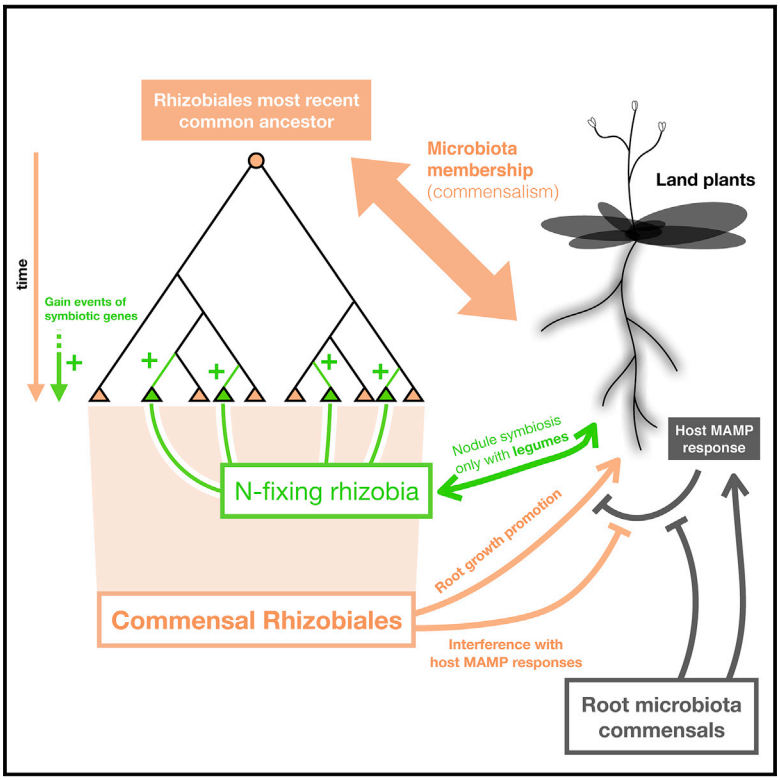
Resource

Cell Host & Microbe

Modular Traits of the Rhizobiales Root Microbiota and Their Evolutionary Relationship with Symbiotic Rhizobia

Graphical Abstract



Highlights

- Comparative genomics of >1,300 Rhizobiales isolated from diverse plant hosts
- Commensalism predates nitrogen-fixing nodule symbiosis in rhizobia
- Root growth promotion is a conserved trait of commensal Rhizobiales
- Microbiota members exhibit strain-specific interference of *Arabidopsis* immune responses

Authors

Ruben Garrido-Oter, Ryohei Thomas Nakano, Nina Dombrowski, Ka-Wai Ma, The AgBiome Team, Alice C. McHardy, Paul Schulze-Lefert

Correspondence

schlef@mpipz.mpg.de

In Brief

Garrido-Oter, Nakano, Dombrowski, et al. employ comparative genomics of ~1,000 Rhizobiales strains that were isolated from diverse plant species to reconstruct the evolutionary history of nitrogen-fixing symbiosis. *Arabidopsis* re-colonization experiments reveal that root growth promotion and interference with host immune responses are modular traits of the root microbiota.

Garrido-Oter et al., 2018, Cell Host & Microbe *24*, 155–167
July 11, 2018
https://doi.org/10.1016/j.chom.2018.06.006

CellPress

Cell Host & Microbe
Resource

Modular Traits of the Rhizobiales Root Microbiota and Their Evolutionary Relationship with Symbiotic Rhizobia

Ruben Garrido-Oter,[1,2,6] Ryohei Thomas Nakano,[1,2,6] Nina Dombrowski,[1,3,6] Ka-Wai Ma,[1] The AgBiome Team,[4] Alice C. McHardy,[5] and Paul Schulze-Lefert[1,2,7,*]
[1]Department of Plant Microbe Interactions, Max Planck Institute for Plant Breeding Research, Cologne 50829, Germany
[2]Cluster of Excellence on Plant Sciences, Dusseldorf 40225, Germany
[3]University of Texas Austin, Marine Science Institute, Port Aransas, TX 78373, USA
[4]AgBiome, 104 T.W. Alexander Drive, Building 1, Research Triangle Park, NC 27709, USA
[5]Department of Computational Biology of Infection Research, Helmholtz Center for Infection Research, Braunschweig 38124, Germany
[6]These authors contributed equally
[7]Lead Contact
*Correspondence: schlef@mpipz.mpg.de
https://doi.org/10.1016/j.chom.2018.06.006

SUMMARY

Rhizobia are a paraphyletic group of soil-borne bacteria that induce nodule organogenesis in legume roots and fix atmospheric nitrogen for plant growth. In non-leguminous plants, species from the Rhizobiales order define a core lineage of the plant microbiota, suggesting additional functional interactions with plant hosts. In this work, genome analyses of 1,314 Rhizobiales isolates along with amplicon studies of the root microbiota reveal the evolutionary history of nitrogen-fixing symbiosis in this bacterial order. Key symbiosis genes were acquired multiple times, and the most recent common ancestor could colonize roots of a broad host range. In addition, root growth promotion is a characteristic trait of Rhizobiales in *Arabidopsis thaliana*, whereas interference with plant immunity constitutes a separate, strain-specific phenotype of root commensal Alphaproteobacteria. Additional studies with a tripartite gnotobiotic plant system reveal that these traits operate in a modular fashion and thus might be relevant to microbial homeostasis in healthy roots.

INTRODUCTION

Extracellular non-self perception in plants is typically mediated by plasma membrane-bound pattern recognition receptors (PRRs) that detect conserved microbial epitopes, termed microbe-associated molecular patterns (MAMPs) (Boller and Felix, 2009). Activation of a PRR by a cognate MAMP induces plant defense responses, collectively termed MAMP-triggered immunity (MTI), which limit pathogen growth (Jones and Dangl, 2006). While early MTI responses (<1 hr), such as reactive oxygen species burst and extensive transcriptional and metabolic reprogramming, are necessary to limit bacterial growth (Li et al., 2014), chronic exposure of seedlings to the bacterial flagellin peptide flg22 results in MTI-mediated inhibition of shoot and root growth (Gómez-Gómez et al., 1999), which is thought to be an indirect consequence of immune signaling and to reflect growth-defense tradeoffs due to limited resources (Huot et al., 2014). Genes encoding MAMP epitopes, such as flg22, are evolutionarily conserved in bacteria with different lifestyles, but whether non-pathogenic bacteria can also elicit MTI responses and influence growth-defense tradeoffs is unclear.

In nature, healthy plants live in intimate association with diverse bacteria, called the plant microbiota, which form taxonomically structured communities on above- and belowground organs. Despite extensive variation at lower taxonomic ranks, a subset of bacterial lineages, designated the "core root microbiota," are ubiquitous across a wide range of environments and host species. One of the most abundant groups is the bacterial order Rhizobiales (Hacquard et al., 2015; Yeoh et al., 2017), which includes species known as rhizobia that engage in beneficial interactions with legumes, and are able to induce nodule organogenesis in legume roots and fix atmospheric nitrogen for plant growth. These symbiotic relationships require dedicated signaling molecules (e.g., bacterial Nod factors and root-secreted isoflavonoids; Fisher and Long, 1992; Oldroyd, 2013) that mediate recognition of compatible symbionts to initiate nodule organogenesis, bacterial accommodation, and subsequent transcriptional and metabolic reprogramming in both partners (Colebatch et al., 2004; Oldroyd et al., 2011; Udvardi and Poole, 2013). Interestingly, rhizobia are also able to suppress legume immunity during establishment of symbiosis (Kawaharada et al., 2015; Liang et al., 2013; Okazaki et al., 2013). This raises the questions of whether Rhizobiales that associate with the roots of non-legumes also possess the genetic toolkit required for symbiosis and if they are capable of interfering with root development and immune responses.

To date, isolation and whole-genome sequencing efforts targeting the Rhizobiales order have largely focused on strains

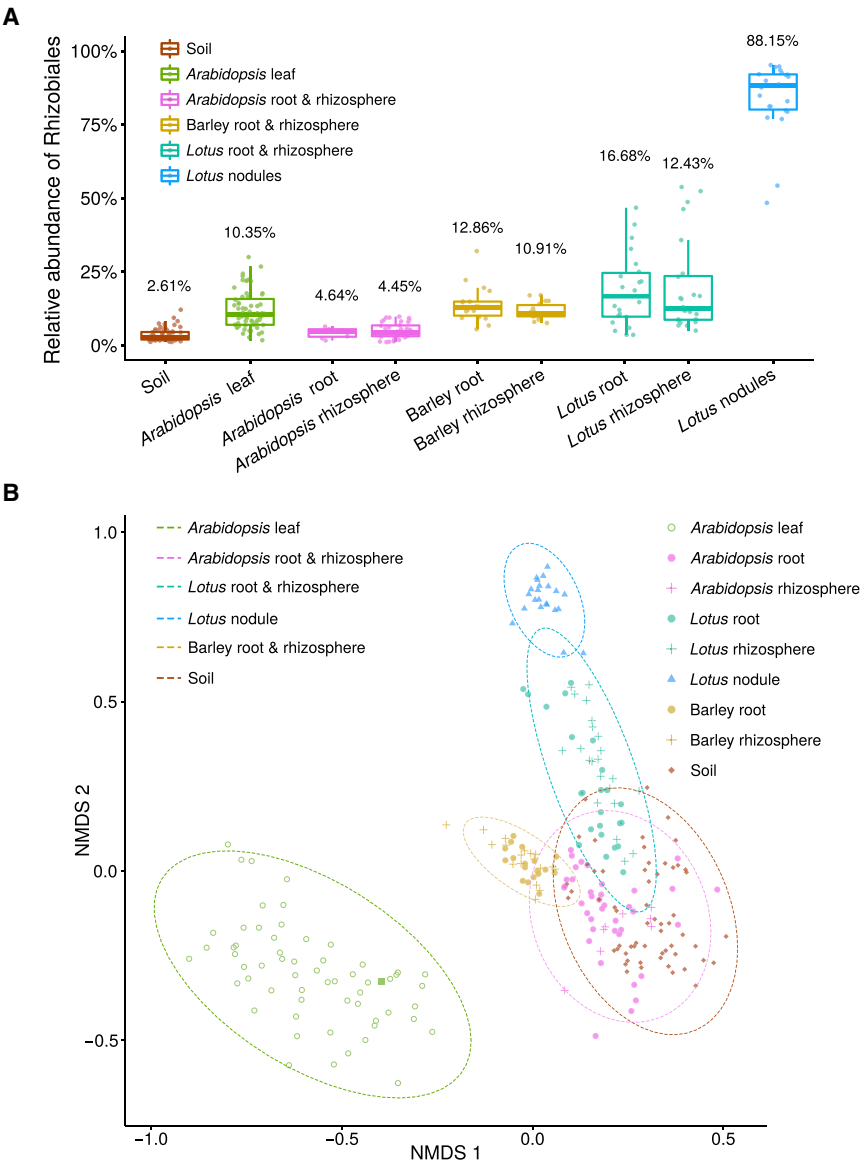


Figure 1. Conserved Microbiota Membership of the Bacterial Order Rhizobiales

Analysis of Rhizobiales community diversity using data from five previous 16S rRNA gene amplicon surveys covering root, rhizosphere, and nodule samples of a taxonomically diverse panel of plant hosts grown in a variety of natural and agricultural soils.

(A) Aggregated relative abundances of Rhizobiales in each host and compartment (n = 453). Percentages show the average contribution of Rhizobiales to each subset of samples.

(B) Analysis of beta-diversity of Rhizobiales between samples across hosts and compartments (indicated by different colors and shapes, respectively). Dashed lines correspond to a Gaussian distribution fitted to each cluster (95% confidence interval). See also Figures S1 and S2 and Tables S1 and S2.

derived from legume nodules, which are almost exclusively colonized by nitrogen-fixing symbionts. Based on these datasets, a common origin and repeated subsequent losses of symbiotic capabilities have been proposed (Raymond et al., 2004). Whether this model holds true for Rhizobiales species enriched in niches other than legume nodules is unknown. We describe here a collection of Rhizobiales isolates from plant- and soil-associated environments of unprecedented size and taxonomic diversity. We compared their whole-genome sequences with those belonging to legume symbionts and reconstructed the evolutionary history of nitrogen-fixing symbiosis in this bacterial order. In addition, large-scale binary and ternary co-inoculation experiments of germ-free *Arabidopsis thaliana* plants and representative native root commensals from Rhizobiales and its sister lineages reveals modular traits and emergent properties of the root microbiota.

RESULTS

Conserved Microbiota Membership of the Bacterial Order Rhizobiales

Large-scale bacterial isolation and DNA sequencing allowed us to assemble high-quality whole-genome sequences of 944 bacterial representatives of the Rhizobiales order. These strains originated from a panel of taxonomically distant plant hosts, including non-legumes, as well as associated environmental sources such as soil, insects, and nematodes (Table S1). These genomes cover known Rhizobiales lineages as well as several lineages previously not present in public databases. We compared these sequences with 370 high-quality genome assemblies of rhizobia, mostly from legume nodule symbionts, retrieved from public databases (Table S1).

We cross-referenced 16S rRNA gene sequences from these genomes with five previous culture-independent amplicon studies of the plant microbiota from five associated micro-habitats (soil, root, rhizosphere, leaf, and legume nodule) of a taxonomically broad set of plant hosts grown in six different soil types (Table S2) (Schlaeppi et al., 2014; Bulgarelli et al., 2015; Bai et al., 2015; Zgadzaj et al., 2016). This high-resolution analysis (99% sequence identity) confirmed that Rhizobiales are consistently found in high relative abundances and enriched in the root and leaf communities of phylogenetically diverse plant hosts, and are therefore part of the core plant microbiota (5%–17% mean relative abundance; Figure 1A). Beta-diversity analyses within the order Rhizobiales showed no clear differences between samples obtained from the root, rhizosphere, or unplanted soil (Figure 1B), indicating a conserved adaptation of soil-borne Rhizobiales to the root environment. In contrast, we observed separation between samples from root compartments and those

from leaves and legume nodules, explained by the predominance of *Methylobacterium* in the phyllosphere and the compatible symbiont *Mesorhizobium loti* in *L. japonicus* nodules, respectively (Figure S1).

Evolutionarily Conserved Adaptation of Rhizobiales Sublineages to the Plant Niche

To study the functional diversity of Rhizobiales, we performed a comparative analysis of the 1,314 whole genomes present in the dataset. Rhizobial genomes are typically mosaic and multipartite in structure due to the presence of large extra-chromosomal replicons, which result in large pan-genomes consisting of a stable and conserved core and a variable accessory gene repertoire (Galibert et al., 2001; Young et al., 2006). Permutation-based analysis of annotated and *de novo* predicted gene families revealed a small core (27.30% of genes per isolate; 802 gene families on average) and a vast and largely uncharacterized pan-genome (>100,000 orthologous genes; Figures S2C and S2D). Principal coordinates analysis of genome-wide functional distances based on a calculation of presence and absence profiles of Kyoto Encyclopedia of Genes and Genomes-annotated genes showed a clear differentiation by taxonomic lineage but no separation between nodulating and non-nodulating strains or between root- and soil-derived isolates (Figure S2A). This indicates that the acquisition of symbiosis genes required for interactions with legumes was not accompanied by whole-genome level signatures of adaptation. Furthermore, we found no segregation between root commensal genomes according to host species (Figure S2B), suggesting a broad host range of Rhizobiales microbiota members. Interestingly, a subset of strains isolated from insects showed high whole-genome level similarity to other members of the Rhizobiaceae family (black dots; Figure S2B and Table S1), indicating that insects might act as transmission vectors for rhizobia.

Convergent Evolution of Nitrogen-Fixing Symbiosis in Rhizobiales

To explore the evolutionary history of nitrogen-fixing symbiosis in rhizobia, we performed a phylogenomic reconstruction of ancestral characters for all sequenced genomes. First, a rooted species tree was generated from a multiple sequence alignment of conserved, single-copy, and vertically inherited genes. Next, we reconstructed the presence or absence of key symbiotic genes, the nitrogenase-encoding *nif* and nodulation factor *nod* genes, in each of the ancestral genomes along the branches of the species tree using a maximum likelihood approach (Figures 2 and S3). We identified a high correlation of gain and loss patterns of the corresponding gene families required for legume symbiosis (Figure S2E), indicating concerted acquisition of these genes, likely via horizontal gene transfer. Given this pattern of evolutionarily concurrent gain and loss, we then employed the conserved *nifH* marker as a proxy for symbiotic capability. This revealed that the most recent common ancestor of all rhizobia likely lacked the symbiosis genetic toolkit and that the capacity to form nodules and fix atmospheric nitrogen was acquired multiple independent times after the speciation events that led to the formation of the major rhizobial taxonomic groups (green branches in Figure 2). Phylogenetic analysis of *nifH* sequences found in the current dataset (n = 296; Table S3) and in non-Rhizobiales genomes (n = 585) confirmed that the *nifH* genes found in rhizobial strains belonging to the Bradyrhizobiaceae and Rhizobiaceae families constitute two separate but closely related clades (Figure S2F), suggesting that one or two primordial acquisitions from unknown non-Rhizobiales donor(s) was followed by multiple horizontal gene transfers within this bacterial order.

Conserved Rhizobiales Root Growth Promotion Activity in *A. thaliana*

The enrichment of all major Rhizobiales taxonomic groups in samples from plants grown in natural soils (Figure 1) and the lack of nodulation and nitrogen fixation capability in the majority of isolates from roots (99.74% of strains; Figure 2) predict an ancestral mechanism of interaction with both non-leguminous and leguminous hosts that enables successful root colonization. We tested whether this association influenced host physiology by performing binary interaction experiments with a panel of Rhizobiales isolates and germ-free *A. thaliana* plants co-cultivated under laboratory conditions. Guided by the phylogenomic reconstruction, we selected a subset of taxonomically and functionally diverse strains belonging to all major Rhizobiales clades of the *A. thaliana* root microbiota, as well as characterized isolates for nitrogen-fixing nodule symbiosis (rhizobia) or for transfer DNA delivery into plant cells (*Agrobacterium*) (arrowheads in Figure 2). We grew wild-type *A. thaliana* plants for 3 weeks on agarose media containing individual bacterial strains and measured primary root length and shoot fresh weight. The majority of the tested strains elicited robust root growth promotion (RGP) (Figure 3; Table S4). Importantly, this phenotype was detected in isolates belonging to the five major taxonomic groups that contain nitrogen-fixing nodule symbionts and *A. thaliana* root-derived commensals, suggesting that this trait is characteristic of the Rhizobiales order. All isolates that do not affect root growth were undetectable in roots except for NGR234 (isolates 157, 635, and 685; Figure S4A), indicating that the lack of the RGP phenotype is mostly due to their inability to colonize the plant host in our experimental setup. Of note, representative *A. thaliana* root-derived commensals from sister core root microbiota lineages within the Alphaproteobacteria (Caulobacterales and Sphingomonadales orders) showed no RGP phenotype despite efficient root colonization (Figures 3 and S4A). Experiments with four *A. thaliana* accessions and multiple exemplars of the most abundant Rhizobiales species found in the *A. thaliana* root microbiota (Bulgarelli et al., 2012) indicated that RGP is a conserved trait in host and microbial populations (Figures S5A and S5B). Furthermore, binary interaction experiments with *A. thaliana* grown under nutrient-limiting conditions revealed that Rhizobiales isolates consistently rescued phosphate starvation-induced root growth inhibition (RGI) and retained RGP under nitrogen starvation (Figures S5C–S5E). This suggests that, unlike nodule symbiosis (Carroll and Gresshoff, 1983), commensal RGP in *A. thaliana* operates both under nitrogen-replete and -deficient conditions.

Interference with Root Meristem Homeostasis and MAMP-Triggered Responses

To gain deeper insights into how Rhizobiales commensals influence root development, we performed time-resolved

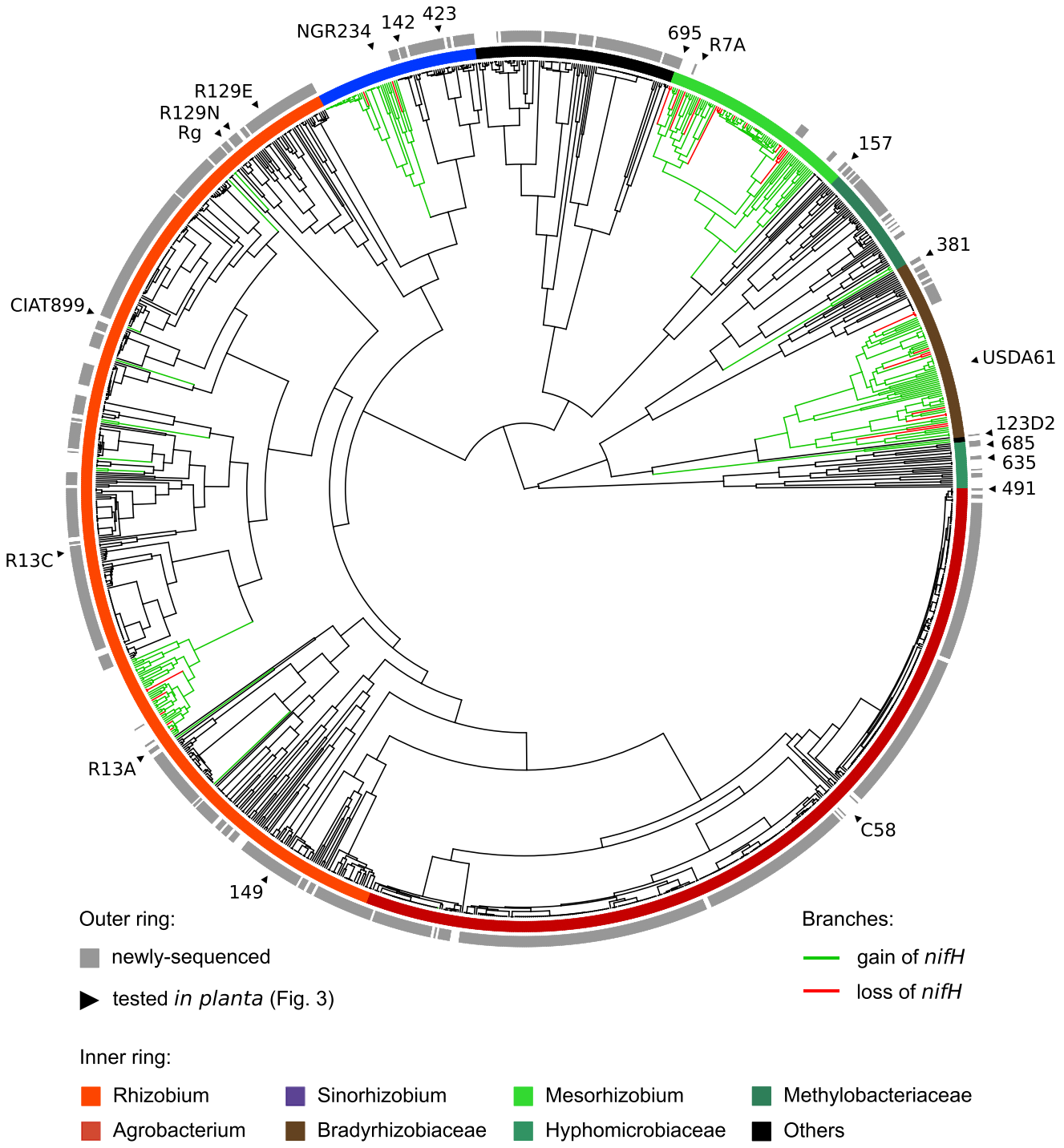


Figure 2. Convergent Evolution of Nitrogen-Fixing Symbiosis in Rhizobiales

Phylogenetic tree of rhizobia and maximum likelihood reconstruction of ancestral symbiotic genotypes. Phylogenetic tree of sequenced isolates (n = 1,314) inferred from aligned single-copy marker genes using a Bayesian approach. Different taxonomic groups are indicated by the various colors in the first ring. The second ring depicts newly sequenced isolates in gray (n = 944). Green branches correspond to likely gains of symbiosis genes whereas those in red correspond to probable losses. See also Figures S2 and S3 and Table S1.

developmental and transcriptomic analyses with *Rhizobium* 129E that lacks *nif* and *nod* genes (Figures 2 and S3) and belongs to one of the most abundant operational taxonomic units enriched in the *A. thaliana* root microbiota (Bulgarelli et al., 2012). Binary interaction experiments with wild-type *A. thaliana* Col-0 plants revealed an enlargement of the root meristematic

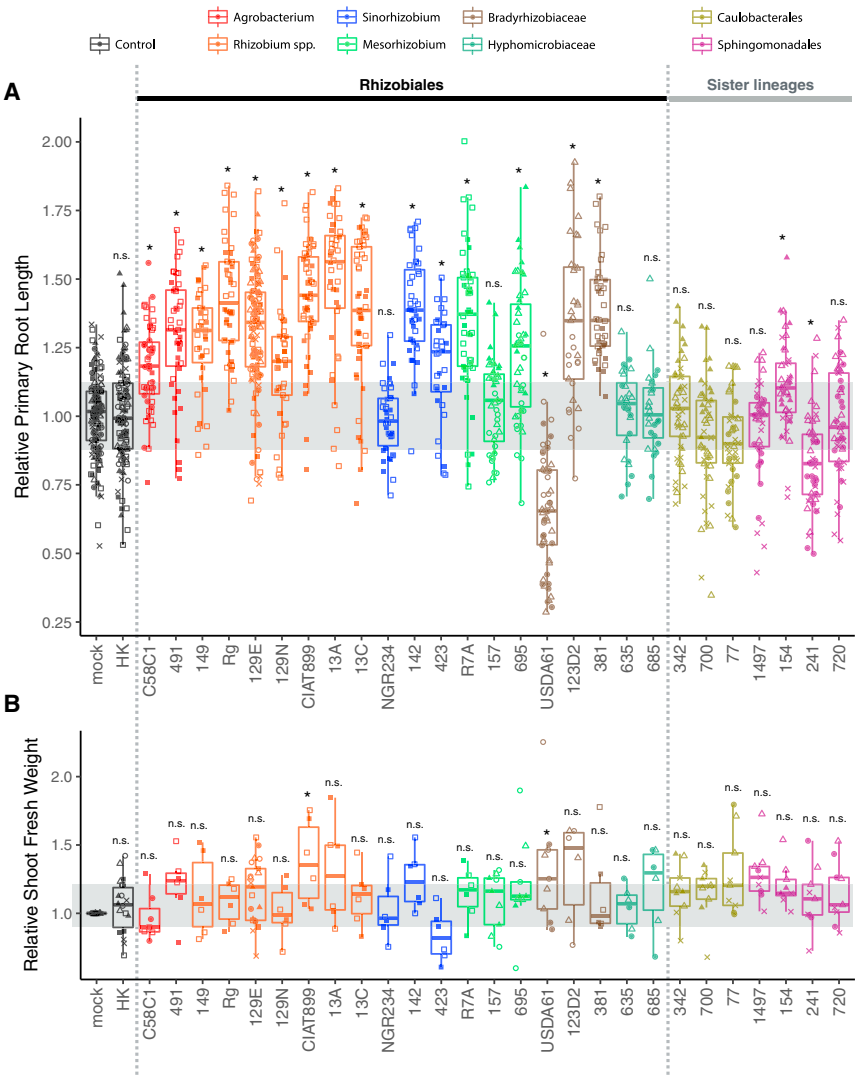


Figure 3. Conserved Rhizobiales Root Growth Promotion Activity in *Arabidopsis thaliana*

Binary interaction experiments between isolates from Rhizobiales or sister lineages and germ-free *A. thaliana* wild-type plants (Col-0) grown on agarose media. Macroscopic plant phenotypes after co-cultivation on agarose media containing individual isolates were recorded after 3 weeks.

(A) Primary root length relative to mock control for each treatment as well as an additional heat-killed bacteria control (n = 1,487).

(B) Shoot fresh weight relative to mock control for each treatment and a heat-killed bacteria control (n = 247). Different shapes depict data points from seven separate full-factorial replicates. Asterisks indicate statistical significance corresponding to a Dunnett's test with false discovery rate correction (α = 0.05). See also Figures S4 and S5 and Table S4.

zone at 14 days post inoculation (dpi) (approximately 2-fold; Figure 4), whereas cell growth in root elongation and differentiation zones remained indistinguishable from germ-free control plants (Figures S6A–S6D). While antagonism between the phytohormones cytokinin and auxin is critical for root meristem homeostasis (Dello Ioio et al., 2008), RGP was retained in *A. thaliana* mutants lacking a cytokinin receptor (*ahk3-3*) or cytokinin signaling components (*arr1arr12*; Figures S6E and S6F), as well as in mutants lacking either an auxin receptor (*tir1-1*), auxin transporters (*aux1-7* and *axr4*), or auxin signaling components (*arf7arf19*; Figure S6G).

To further investigate the mechanism(s) by which *Rhizobium* 129E interferes with host physiology, we performed RNA sequencing (RNA-seq) of *A. thaliana* roots after inoculation with *Rhizobium* 129E. To identify plant genes responding to bacterial colonization independently of nutrient status, we collected samples at four time points (4, 8, 12, and 16 dpi) under P_i-sufficient and P_i-deficient conditions, both of which resulted in commensal-mediated RGP (Figure 5A). Bacterial root colonization under both nutrient conditions was comparable and reached a plateau at 12 dpi after multiplication by three $\log_{10}$ units (Figure S4B). We found 1,771 genes to be differentially expressed in inoculated roots with respect to mock controls for at least one time point, demonstrating that the interaction with *Rhizobium* 129E results in dynamic host transcriptional reprogramming (Figure 5B; Table S5). Strikingly, four gene clusters (Clusters 1, 15, 13, and 6; 433 genes) were significantly enriched in functional categories related to immune responses and mostly downregulated by colonization of *Rhizobium* 129E (Figure 5C). This finding prompted us to compare our sequencing data with a published dataset of transcriptional changes in roots in response to flg22 treatment (Castrillo et al., 2017). We observed a significant anti-correlation between the transcriptional response to root colonization by *Rhizobium* 129E and treatment with bacterial pathogen-derived flg22 (Figure 5D), implying that this commensal can partially suppress MAMP-triggered transcriptional reprogramming.

A long-known output of MTI is RGI (Gómez-Gómez et al., 1999), which is believed to reflect a tradeoff between growth and defense in plants (Huot et al., 2014; Lozano-Durán and Zipfel, 2015). We examined whether *Rhizobium* 129E can attenuate RGI upon chronic exposure of roots to flg22 and found that the commensal can fully override RGI (Figure 6A), further supporting its capacity to suppress MAMP-induced responses. We explored the prevalence of this trait among microbiota strains belonging to Alphaproteobacteria (Figure 6B). None out of four further tested Rhizobiales strains (isolates 491, 13A, 142, and 423) overcame flg22-stimulated RGI, demonstrating strain-specificity of this trait and that the RGP character and MAMP response suppression are separable traits among root commensal Rhizobiales. Sphingomonadales 1497 fully

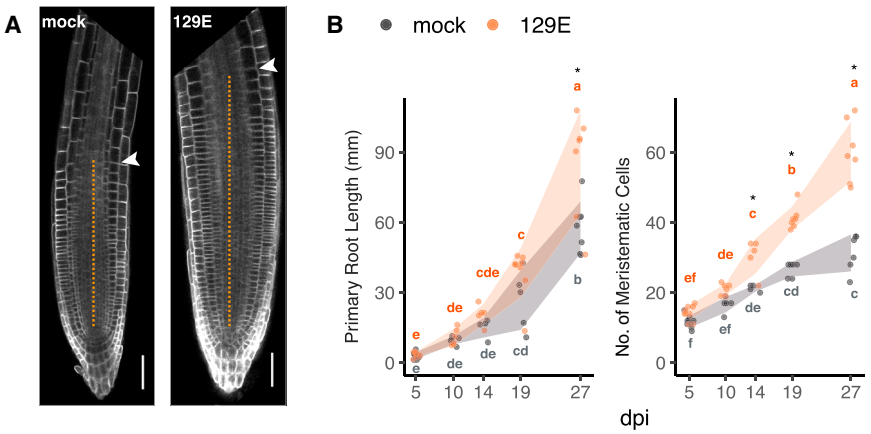


Figure 4. Interference with Root Meristem Homeostasis by *Rhizobium* 129E

(A) Confocal micrographs in the meristematic zone (MZ) (indicated by dotted lines) of the roots of transgenic *A. thaliana* (Wave_131Y) expressing a plasma membrane-targeted yellow fluorescent protein inoculated with 129E or with a mock control. Arrowheads indicate the place of transition from cell division to cell elongation. Bars correspond to 50 μm.

(B) Number of cells within the MZ was recorded with primary root length at 5, 10, 14, 19, and 27 days post inoculation (n = 65). Shaded areas indicate means ±SD. Letters indicate statistical significance corresponding to Tukey's HSD test (α = 0.05) and asterisks highlight significant difference between mock and *Rhizobium* 129E-treated roots within each time point. See also Figure S6 and Table S4.

suppressed flg22-induced RGI despite a lack of RGP activity, revealing that suppression of MAMP-induced responses also evolved in another taxonomic lineage of the core root microbiota. The RGP activity of all non-suppressive Rhizobiales isolates was maintained in the mutant lacking the flg22 receptor FLAGELLIN SENSITIVE2 (FLS2; Gómez-Gómez et al., 1999) in the presence of this MAMP epitope, indicating that loss of bacterial RGP activity upon flg22 treatment is caused by MAMP-stimulated plant responses. Collectively, these data suggest that *Rhizobium* 129E and Sphingomonadales 1497 commensals can influence the outcome of defense-growth tradeoffs in the presence of MAMPs. To test the potential relevance of MAMP response suppression in a microbiota context, we analyzed the genome sequences of the taxonomically representative *At*-SPHERE culture collection of the root microbiota derived from *A. thaliana* roots (Bai et al., 2015). This revealed widespread presence of protein sequences containing potential MAMP epitopes (flg22, flgII-28, elf18, csp22, and nlp20; Böhm et al., 2014; Cai et al., 2011; Felix and Boller, 2003; Felix et al., 1999; Kunze et al., 2004) in bacteria isolated from healthy plants (Figure S7A). Consistently, three out of four tested root-derived Pseudomonadales from the *At*-SPHERE collection activated MTI reporter gene expression in the roots of transgenic *A. thaliana* $CYP71A12_{pro}$*:GUS* seedlings (Figure S7B) (Millet et al., 2010).

Modularity of Commensal RGP and Interference with MAMP Responses

The phenotypic variation observed in mono-associations with root commensals isolated from *A. thaliana* grown in the same soil raised the question of whether RGP and suppression of MAMP-induced outputs operate independently in a microbial community context. We tested for potential functional modularity of these bacterial traits by selecting commensals with distinct combinations of these traits from the Rhizobiales order (isolates 129E and 142) as well as from sister clades of Alphaproteobacteria (isolates 1497 and 700) for tripartite co-inoculation experiments (Figure 7). We measured primary root growth of *A. thaliana* after co-inoculation with all single and pairwise combinations of bacterial isolates in the presence or absence of exogenous flg22. This revealed that RGP and MAMP response suppression are largely dominant traits, expressed whenever one of the bacterial partners has the respective function (Figure 7). For example, co-inoculation of the root growth-promoting and MAMP response-suppressive isolate *Rhizobium* 129E with Caulobacterales 700, which lacks either character, resulted in retention of both phenotypes. Importantly, co-inoculation of *Sinorhizobium* 142, which promotes root growth only in the absence of flg22, with the MAMP response-suppressive strain Sphingomonadales 1497, which lacks RGP activity, conferred a full RGP phenotype under both mock and flg22 treatments (Figure 7). In this combination, RGP in the presence of flg22 cannot be explained by the sum of phenotypes seen in the corresponding mono-associations because none of the strains showed RGP under this condition, suggesting that these modular traits operate synergistically and are an emergent property of this tripartite interaction.

DISCUSSION

Adaptation to the Root Environment Predates Acquisition of Symbiosis Genes

The majority of bacterial species within the Rhizobiales bacterial order are consistently enriched in the roots and leaves of legume and non-legume plant species and are therefore core members of the plant microbiota (Figure 1A). The high similarity between the Rhizobiales communities in soil, root, and rhizosphere samples collected from diverse plant species grown in six different soil types suggests that the majority of soil-borne Rhizobiales are also adapted to the root environment. The prevalence of root-competent Rhizobiales in the soil biome could be explained by a positive plant-soil feedback sustained over evolutionary timescales (Bever et al., 2012), a process whereby plants alter the biotic qualities of the soil in which they grow. Large-scale plant growth assays revealed a consistent RGP phenotype present in isolates from all five major Rhizobiales lineages capable of colonizing *A. thaliana* roots (Figure 3A) without a negative impact on shoot fresh weight (Figure 3B), which is indicative of a commensal lifestyle. Thus, most Rhizobiales are capable of actively interacting with the non-leguminous plant and altering their host niche rather than merely colonizing the root. Further experimentation is needed to elucidate whether the genetic

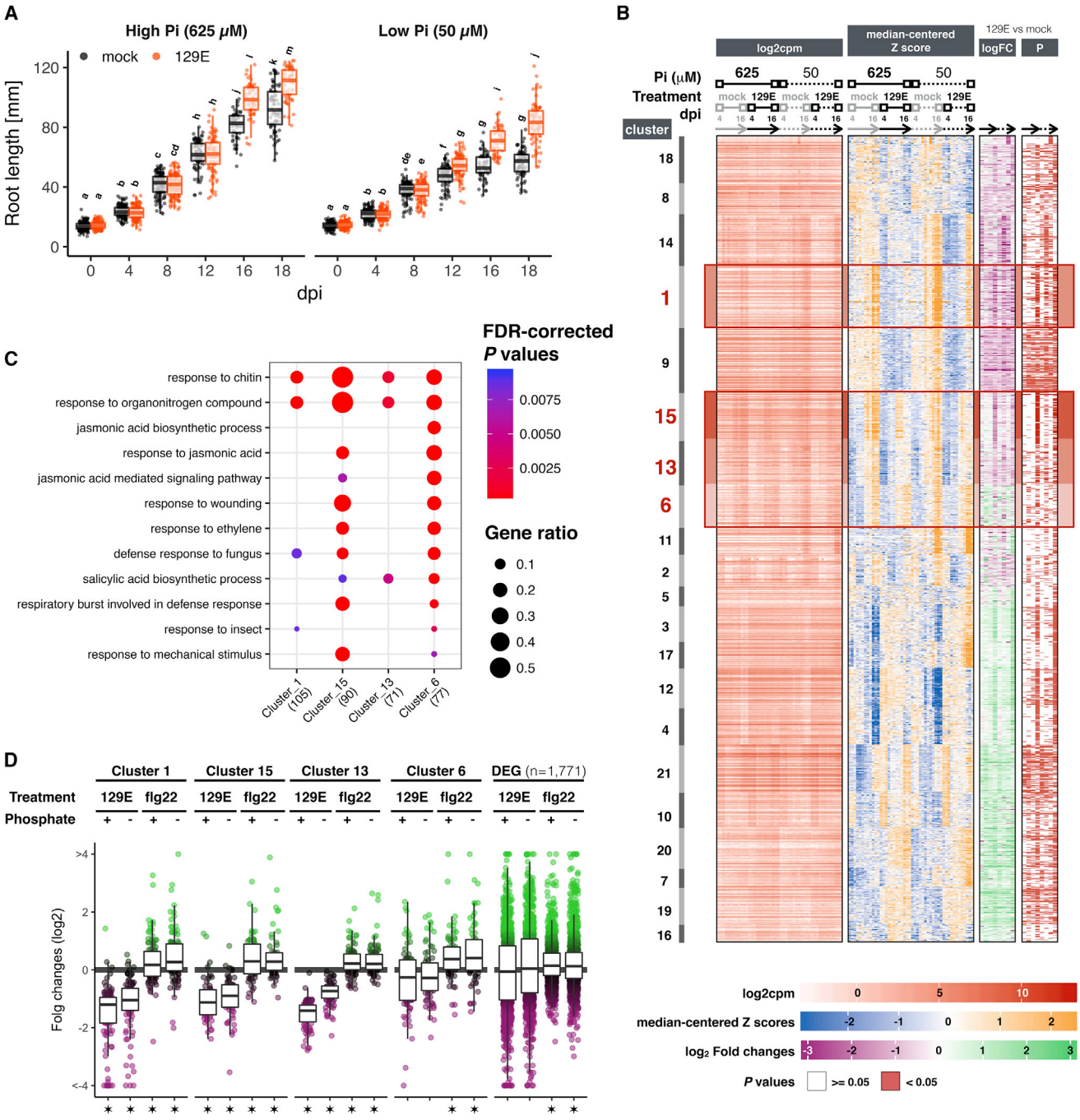


Figure 5. Transcriptional Interference with MAMP-Triggered Responses

(A) Time course measurement of primary root growth under high P_i (625 μM) and low P_i (50 μM) conditions (n = 2,323). Letters indicate statistical significance corresponding to Tukey's HSD test corrected for multiple comparisons ($\alpha = 0.05$).

(B) Heatmaps showing expression level (log_2 counts per million; log2cpm), expression pattern (gene-wise median-centered *Z* scores), response to *Rhizobium* 129E (log_2 fold changes; logFC; green and magenta colors indicate up- and downregulation, respectively) and significance (false discovery rate-corrected p values <0.05). Co-expressed gene clusters (cluster) are defined by *k*-means clustering ($k = 21$).

(C) Enrichment analysis of the clusters 1, 15, 13, and 6 identified defense-related gene ontology categories. Gene numbers used for the analysis are shown in parentheses.

(D) Comparison of plant responses to *Rhizobium* 129E at 16 days post inoculation and to chronic 1 μM flg22 treatment for 12 days (Castrillo et al., 2017). DEG, differentially expressed genes. Asterisks indicate statistical significance corresponding to one sample t test ($\mu = 0$) with Bonferroni's correction for multiple comparisons ($\alpha = 0.05$). See also Figure S4 and Tables S4 and S5.

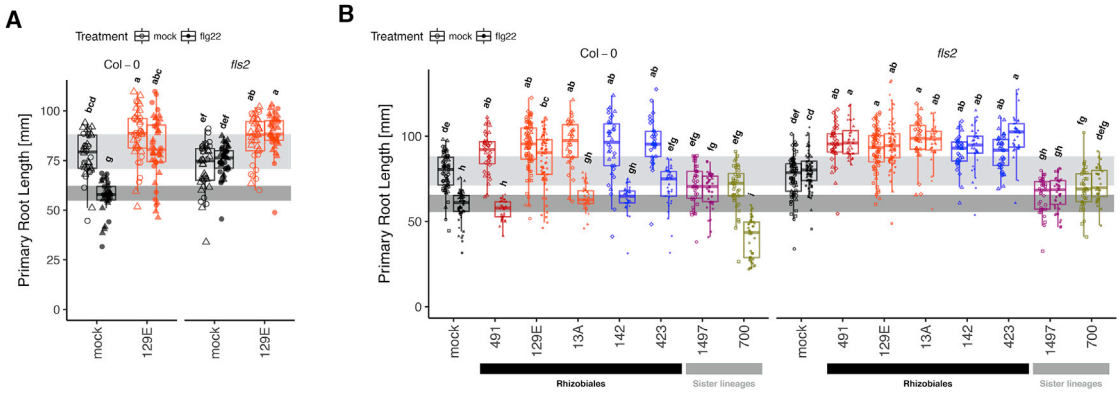


Figure 6. Root Commensals Interfere with MAMP-Triggered Growth Outputs

(A) FLAGELLIN SENSITIVE2 (FLS2)-dependent root growth inhibition induced by chronic exposure to 1 μM flg22 for 14 days was repressed by inoculation with *Rhizobium* 129E (n = 302).

(B) The ability to interfere with flg22-induced root growth inhibition is specific to *Rhizobium* 129E and Sphingomonadales 1497 (n = 1,424). Root growth promotion activity of the other Rhizobiales isolates was abolished in wild-type plants but not in the mutant lacking the flg22 receptor protein FLS2. Open (left boxplots) and closed shapes (right boxplots) within each condition indicate mock and flg22 treatment (1 μM), respectively. Letters indicate statistical significance corresponding to Tukey's HSD test corrected for multiple comparisons ($\alpha = 0.05$). Different shapes depict data points from full-factorial biological replicates. See also Figure S7 and Table S4.

basis and the mechanism of action of this prevalent commensal lifestyle is conserved across Rhizobiales species.

Our dataset, which includes genomes from Rhizobiales strains isolated from root- and soil-associated environmental sources, complements current databases that are highly biased toward genomes of legume nodule symbionts, enabling us to perform a comprehensive analysis of the evolutionary history of the Rhizobiales order and nodule symbiosis. This revealed that the most recent common ancestor of Rhizobiales did not have key symbiosis genes, which were acquired jointly rather than sequentially (Figures 2 and S2E). Thus, robust microbiota membership across flowering plants and a commensal lifestyle clearly predates and possibly predisposed ancestral Rhizobiales to the subsequent acquisition of genes required for nodule symbiosis. A recent study described a symbiont switch from the actinomycete *Frankia* to rhizobia in the non-leguminous *Parasponia* lineage (van Velzen et al., 2018), suggesting that extant rhizobia might have acquired the *nod* and *nif* genes from an outside donor that was already capable of nitrogen-fixing nodule symbiosis such as actinorhizal symbionts. Interestingly, members of the Burkholderiales order, belonging to the class Betaproteobacteria, and representing another core lineage of the root microbiota, have strong plant growth-promoting activity or engage in nodulation and nitrogen-fixing symbiosis with basal legumes (Chen et al., 2003; Masson-Boivin et al., 2009; Moulin et al., 2001; Poupin et al., 2016). This suggests that convergent evolution toward nitrogen-fixing symbiosis from a state of rhizosphere competence and microbiota membership might have relevance beyond Rhizobiales.

Lifestyle Transition from Root Commensalism to Nodule Symbiosis

The capacity of nitrogen-fixing rhizobia to suppress immune responses in legumes has been interpreted as evidence that nodule symbiosis emerged from a pathogenic lifestyle (Cao et al., 2017). Experimental evolution studies that introduced the symbiotic plasmid of the nitrogen-fixing legume symbiont *Cupriavidus taiwanensis* into the root pathogen *Ralstonia solanacearum*, both belonging to the order Burkholderiales, demonstrated a change of lifestyle from saprophytism toward mutualism, accompanied by genome remodeling that enabled efficient engagement with the symbiotic hosts (Capela et al., 2017; Remigi et al., 2016). Our combined phylogenomic (Figures 2 and S2) and phenotypic analyses (Figures 5, 6, and 7), however, suggest another scenario in which an ancestral state of adaptation to the root environment and microbiota membership allowed a shift from commensalism to mutualism. This model predicts that nodulating rhizobia and Rhizobiales root commensals share conserved components for root colonization. A recent genome-wide mutagenesis screen in the alfalfa symbiont *Ensifer meliloti* (formerly *Sinorhizobium meliloti*) revealed that nearly 2% of genes, ~80% of which were chromosomally located, were relevant for rhizosphere colonization (Salas et al., 2017). We tested whether these rhizosphere competence genes are present in other Rhizobiales lineages and non-leguminous root commensals lacking the key symbiotic toolkit and found that, on average, 86% are conserved in each of the 1,314 genomes (Figure S2G). This strongly supports our hypothesis that the most recent ancestor of the Rhizobiales was already fully adapted to the root environment. Moreover, the observation that six extant members of nitrogen-fixing sublineages of rhizobia that were isolated from *A. thaliana* roots (*Sinorhizobium*/*Ensifer*, *Mesorhizobium*, and *Bradyrhizobium* spp.) lost key symbiotic genes suggests that a lifestyle transition from mutualism back to commensalism can also occur. In addition, a recent genome-wide comparative analysis of 37 plant species identified multiple independent loss-of-function events of the crucial symbiotic regulator NODULE INCEPTION (Griesmann et al., 2018), further indicating that selective pressures against

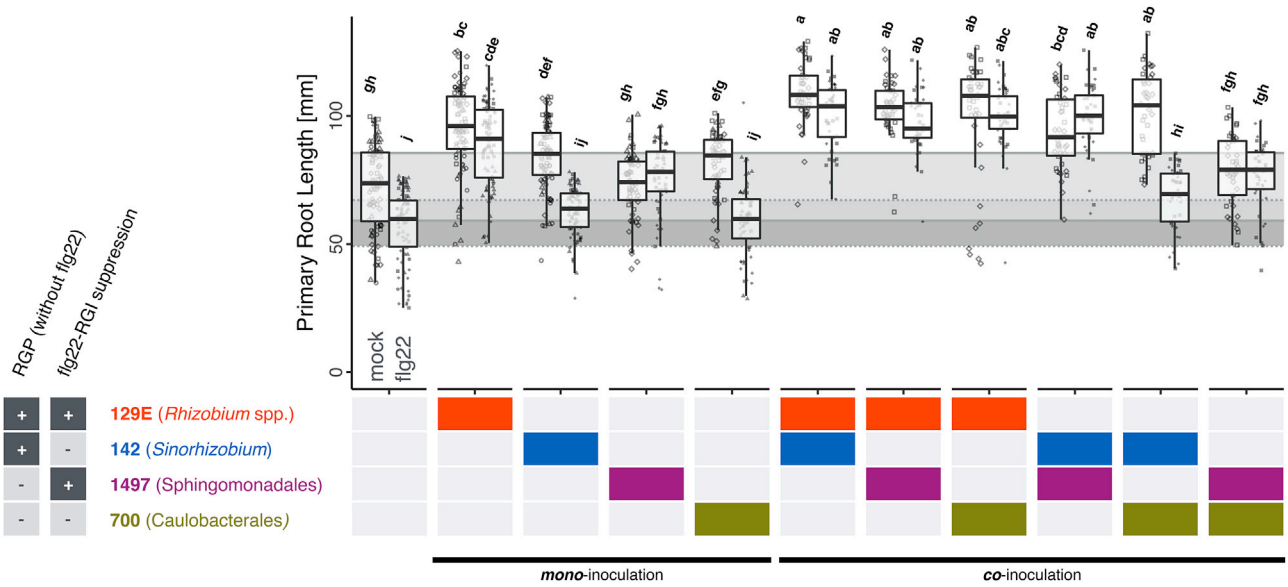


Figure 7. Modularity of Commensal Root Growth Promotion and Interference with MAMP Responses

Primary root growth under binary and tripartite inoculations was measured in the presence and absence of 1 μM flg22 (n = 1,219). Functional profiles of tested isolates are summarized on the left. Individual isolates or combinations of isolates were mono- or co-inoculated with germ-free *A. thaliana*. Inoculated isolates are depicted in the bottom table with colors. Left and right boxplots within each condition with open and closed shapes indicate mock and flg22 treatments, respectively. Different shapes depict data points from full-factorial biological replicates. Letters indicate statistical significance corresponding to Tukey's HSD test corrected for multiple comparisons ($\alpha = 0.05$). See also Figure S7 and Table S4.

nitrogen-fixing nodule symbiosis can act on both the host and the symbiont.

Diversification of Bacterial Interference with Root Development

Our finding that commensal Rhizobiales are able to manipulate root development is consistent with recent studies showing altered *A. thaliana* root architecture in association with nitrogen-fixing *Mesorhizobium loti* MAFF303099 or *Rhizobium* IRBG74 (Poitout et al., 2017; Zhao et al., 2017). However, whereas in those studies RGI was reported, similarly to *Bradyrhizobium elkanii* USDA61 in our survey, we observed an increase in primary root length with all other root-colonizing isolates (Figure 3). We favor the view that these differences indicate *bona fide* mechanistic diversity in the manipulation of root development. First, while interference with root development by *Rhizobium* 129E is independent of the auxin receptor TIR1 and transporter AUX1, these proteins play a major role in the RGI activity of nodule symbionts *M. loti* MAFF303099 and *Rhizobium* IRBG74 in *A. thaliana*. Second, the activation of cytokinin signaling pathways is essential for nodule organogenesis in legumes (via *Lj*LHK1 or *Mt*CRE1 receptors; Oldroyd et al., 2011), but reduces root meristem size in *A. thaliana* (Dello Ioio et al., 2007). Finally, a well-characterized plant growth-promoting rhizobacterium (PGPR), *Pseudomonas simiae* WCS417, also increases root meristem size, although this bacterium confers RGI by suppressing cell elongation (Zamioudis et al., 2013). Overall, these findings suggest diverse machineries underlying the alterations of host root architecture mediated by symbiotic rhizobia, PGPR *Pseudomonas*, and commensal Rhizobiales.

Commensal Rhizobiales Interfere with MAMP-Induced Responses in *A. thaliana* Independently of Nod Factor Production

Comparison of our time-resolved transcriptomic data with published datasets of flg22-treated *A. thaliana* roots showed that the majority of genes induced by MAMP perception were downregulated in response to *Rhizobium* 129E colonization (Figure 5D), suggesting the ability to interfere with MAMP-induced host transcriptional responses in the absence of external MAMPs. *Rhizobium* 129E, however, lacks the bacterial type III secretion system components and Nod factor biosynthesis genes, which are necessary for MTI suppression by phytopathogenic bacteria (Alfano and Collmer, 2004) or legume nodule symbionts (Liang et al., 2013), respectively. This implies that this commensal strain interferes with host MAMP-induced transcriptional responses independently from previously described MTI suppression pathways. *In planta* proliferation of *Rhizobium* 129E reached a plateau at 12 dpi after exponential growth by ~3 $\log_{10}$ units without recognizable detrimental effects on the host (Figures 5A and S4B), suggesting that the plant can still control bacterial proliferation despite the suppressed MAMP-induced transcriptional response in the host roots.

Unlike RGP, representative Rhizobiales showed population-level diversity and strain-specific ability to override RGI in *A. thaliana* in the presence of flg22 (Figure 6). This indicates that Rhizobiales RGP and interference with flg22-induced RGI are separable traits, excluding the possibility that RGP is merely an indirect consequence of suppressing MAMP-stimulated RGI. This also demonstrates that interference with host MAMP responses by *Rhizobium* 129E (Figures 5 and 6) is due to a direct interaction rather than representing a general plant response to

the presence of bacterial cells. In addition, an isolate from a sister lineage within the Alphaproteobacteria suppressed MAMP-activated RGI without RGP activity (isolate 1497; Figure 6B), showing that other members of the core root microbiota share the ability to interfere with host MAMP responses. This is further supported by previous reports showing that PGPR strains from Gammaproteobacteria (*P. simiae* WCS471; Stringlis et al., 2018) and Firmicutes (*Bacillus subtilis* FB17; Lakshmanan et al., 2013) can suppress *A. thaliana* defense responses. Taken together with our findings, this points to a sparse but taxonomically wide distribution of suppression of MAMP-induced responses in bacterial populations isolated from the roots of healthy plants and suggests that this trait may have been frequently gained and/or lost by microbiota members. Further, this implies that the plant innate immune system can exert selective pressure on root microbiota members to acquire and maintain trait(s) involved in suppressing MAMP-induced responses. The widespread presence of sequences encoding peptides of high similarity to canonical MAMPs in the genomes of the *A. thaliana* root microbiota (Figure S7), and the observed dominant character of the MAMP response interference trait (Figure 6B), indicate that microbial homeostasis in healthy roots may depend on the co-occurrence of commensal community members with MTI-activating and MTI-suppressing activities.

Modular Traits and Emergent Properties of the Plant Microbiota

Co-inoculation experiments of wild-type *A. thaliana* plants with combinations of commensal Alphaproteobacteria isolates with varying capacities for RGP and suppression of flg22-triggered RGI reveal that these traits can act synergistically, resulting in an emergent property of the tripartite system and plant physiological outputs that cannot be predicted by the behavior of the individual strains (RGP of flg22-treated plants; Figure 7). This can be explained by interactions between these two traits, where host MAMP responses disable RGP either by directly interfering with bacterial activity or by reducing host sensitivity. This complex behavior might also explain the phenotypic diversity of RGP and MAMP response suppression among members of the root microbiota and suggests that these bacterial traits impact the fitness of the entire community and the host. We speculate that variation in the efficacy of field applications of PGPR results from similar complex interactions with resident microbiota members on crops (Ansari et al., 2015). Furthermore, despite the reductionism of our experimental setup we observed complex interactions between microbial traits acting in a modular and synergistic manner. Our findings illustrate the necessity of bottom-up experimental approaches, in which populations of microbiota members are first systematically sampled in an unbiased manner, followed by characterization of representative isolates in mono-associations, and finally in a community context under controlled conditions.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Bacterial Strains
  - Plant Model
  - Growing Conditions for Plant Models
  - Culture Conditions for *In Vitro* Systems
- METHOD DETAILS
  - Isolation and Sequencing of Bacterial Strains
  - Bacterial Inoculation and Image Analysis
  - DNA Extraction and Quantitative PCR
  - Confocal Microscopy and Image Analysis
  - RNA Extraction and RNA-seq
  - GUS Histochemical Assay
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Genome Assembly
  - Genome Annotation and Orthology Inference
  - Natural Community Amplicon Sequencing Meta-Analysis
  - Comparative Genomics and Ancestral Character Reconstruction
  - Image Analysis for Root Growth Quantification
  - Image Analysis for Root Developmental Analysis
  - RNA-seq Data Analysis
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and six tables and can be found with this article online at https://doi.org/10.1016/j.chom.2018.06.006.

CONSORTIA

The AgBiome Team: Elka Armstrong, Stacey Badders, Vadim Beilinson, Brooke Bissinger, Courtney Bogard, Mauricio Borgen, Kelly Craig, Larry Daquioag, Billie Espejo, Andrew Graham, Philip E. Hammer, James R. Henriksen, Janice C. Jones, Karen Juan, Rebekah Kelly, Dylan Kraus, Mary Kroner, Larry Nea, Narendra Palekar, Jessica Parks, Jacob Pearce, Vinh D. Pham, Tracy Raines, Kira Bulazel Roberts, Mark Roberts, Steve J. Ronyak, Maureen Schirtzinger, Amy Shekita, Amber Smith, Murray Spruill, Kelly S. Smith, Rebecca Thayer, Dan Tomso, Jake Trimble, Mathias Twizeyimana, Katharine Tyson, Scott Uknes, Sandy Volrath, Eric Ward, Christy Wiggins, Kelly Williamson.

ACKNOWLEDGMENTS

The authors would like to thank Paloma Duran, Dr. Rafal Zgadzaj, Dr. Kenichi Tsuda, Dr. Jane Parker, and Dr. Simona Radutoiu for their useful comments on this manuscript. We also thank Neysan Donnelly and Jack Roach for their comments and help editing the manuscript. This work was supported by funds from the Max Planck Society, a European Research Council advanced grant (ROOTMICROBIOTA), and the "Cluster of Excellence on Plant Sciences" (CEPLAS) program funded by the Deutsche Forschungsgemeinschaft.

AUTHOR CONTRIBUTIONS

R.G.-O., R.T.N., N.D., A.C.M., and P.S.-L. designed the research. The AgBiome Team collected samples, isolated strains, and assembled genomes for 826 bacteria. R.T.N. and N.D. performed and analyzed plant phenotypic experiments. N.D. performed RNA-seq experiments and R.T.N. analyzed those data. R.G.-O. conducted the amplicon data meta-analysis, genome assembly, and annotation and performed comparative genomic analyses and statistical tests. K.-W.M. performed the MTI activation assay. R.G.-O., R.T.N., N.D., K.-W.M., and P.S.-L. interpreted the data, and R.G.-O., R.T.N., and P.S.-L. wrote the manuscript.

DECLARATION OF INTERSTS

Paul Schulze-Lefert is a co-founder of AgBiome and a member of its scientific advisory board.

Received: January 29, 2018
Revised: April 16, 2018
Accepted: June 15, 2018
Published: July 11, 2018

REFERENCES

Alfano, J.R., and Collmer, A. (2004). Type III secretion system effector proteins: double agents in bacterial disease and plant defense. Annu. Rev. Phytopathol. *42*, 385–414.

Ansari, M.F., Tipre, D.R., and Dave, S.R. (2015). Efficiency evaluation of commercial liquid biofertilizers for growth of *Cicer aeritinum* (chickpea) in pot and field study. Biocatal. Agric. Biotechnol. *4*, 17–24.

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

Bever, J.D., Platt, T.G., and Morton, E.R. (2012). Microbial population and community dynamics on plant roots and their feedbacks on plant communities. Annu. Rev. Microbiol. *66*, 265–283.

Böhm, H., Albert, I., Oome, S., Raaymakers, T.M., Van den Ackerveken, G., and Nurnberger, T. (2014). A conserved peptide pattern from a widespread microbial virulence factor triggers pattern-induced immunity in *Arabidopsis*. PLoS Pathog. *10*, e1004491.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Boller, T., and Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annu. Rev. Plant Biol. *60*, 379–406.

Bulgarelli, D., Garrido-Oter, R., Münch, P.C., Weiman, A., Dröge, 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.

Bulgarelli, D., Rott, M., Schlaeppi, K., Ver Loren van Themaat, E., Ahmadinejad, N., Assenza, F., Rauf, P., Huettel, B., Reinhardt, R., Schmelzer, E., et al. (2012). Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. Nature *488*, 91–95.

Cai, R., Lewis, J., Yan, S., Liu, H., Clarke, C.R., Campanile, F., Almeida, N.F., Studholme, D.J., Lindeberg, M., Schneider, D., et al. (2011). The plant pathogen *Pseudomonas syringae* pv. tomato is genetically monomorphic and under strong selection to evade tomato immunity. PLoS Pathog. *7*, e1002130.

Cao, Y., Halane, M.K., Gassmann, W., and Stacey, G. (2017). The role of plant innate immunity in the legume-rhizobium symbiosis. Annu. Rev. Plant Biol. *68*, 535–561.

Capela, D., Marchetti, M., Clerissi, C., Perrier, A., Guetta, D., Gris, C., Valls, M., Jauneau, A., Cruveiller, S., Rocha, E.P.C., et al. (2017). Recruitment of a lineage-specific virulence regulatory pathway promotes intracellular infection by a plant pathogen experimentally evolved into a legume symbiont. Mol. Biol. Evol. *34*, 2503–2521.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods *7*, 335–336.

Carroll, B.J., and Gresshoff, P.M. (1983). Nitrate inhibition of nodulation and nitrogen-fixation in white clover. Z. Pflanzenphysiol. *110*, 77–88.

Castrillo, G., Teixeira, P.J., Paredes, S.H., Law, T.F., de Lorenzo, L., Feltcher, M.E., Finkel, O.M., Breakfield, N.W., Mieczkowski, P., Jones, C.D., et al. (2017). Root microbiota drive direct integration of phosphate stress and immunity. Nature *543*, 513–518.

Chen, W.-M., Moulin, L., Bontemps, C., Vandamme, P., Bena, G., and Boivin-Masson, C. (2003). Legume symbiotic nitrogen fixation by beta-proteobacteria is widespread in nature. J. Bacteriol. *185*, 7266–7272.

Chin, C.-S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E.E., et al. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat. Methods *10*, 563–569.

Colebatch, G., Desbrosses, G., Ott, T., Krusell, L., Montanari, O., Kloska, S., Kopka, J., and Udvardi, M.K. (2004). Global changes in transcription orchestrate metabolic differentiation during symbiotic nitrogen fixation in *Lotus japonicus*. Plant J. *39*, 487–512.

Dello Ioio, R., Linhares, F.S., Scacchi, E., Casamitjana-Martinez, E., Heidstra, R., Costantino, P., and Sabatini, S. (2007). Cytokinins determine *Arabidopsis* root-meristem size by controlling cell differentiation. Curr. Biol. *17*, 678–682.

Dello Ioio, R., Nakamura, K., Moubayidin, L., Perilli, S., Taniguchi, M., Morita, M.T., Aoyama, T., Costantino, P., and Sabatini, S. (2008). A genetic framework for the control of cell division and differentiation in the root meristem. Science *322*, 1380–1384.

Eddy, S.R. (2011). Accelerated profile HMM searches. PLoS Comput. Biol. *7*, e1002195.

Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods *10*, 996–998.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics *27*, 2194–2200.

Emms, D.M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. *16*, 157.

Felix, G., and Boller, T. (2003). Molecular sensing of bacteria in plants. The highly conserved RNA-binding motif RNP-1 of bacterial cold shock proteins is recognized as an elicitor signal in tobacco. J. Biol. Chem. *278*, 6201–6208.

Felix, G., Duran, J.D., Volko, S., and Boller, T. (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. Plant J. *18*, 265–276.

Fisher, R.F., and Long, S.R. (1992). *Rhizobium*–plant signal exchange. Nature *357*, 655–660.

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. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science *293*, 668–672.

Gómez-Gómez, L., Felix, G., and Boller, T. (1999). A single locus determines sensitivity to bacterial flagellin in *Arabidopsis thaliana*. Plant J. *18*, 277–284.

Griesmann, M., Chang, Y., Liu, X., Song, Y., Haberer, G., Crook, M.B., Billault-Penneteau, B., Lauressergues, D., Keller, J., Imanishi, L., et al. (2018). Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. Science. https://doi.org/10.1126/science.aat1743.

Gruber, B.D., Giehl, R.F.H., Friedel, S., and von Wirén, N. (2013). Plasticity of the *Arabidopsis* root system under nutrient deficiencies. Plant Physiol. *163*, 161–179.

Hacquard, S., Garrido-Oter, R., Gonzalez, A., Spaepen, S., Ackermann, G., Lebeis, S., McHardy, A.C., Dangl, J.L., Knight, R., Ley, R., et al. (2015). Microbiota and host nutrition across plant and animal kingdoms. Cell Host Microbe *17*, 603–616.

Higuchi, M., Pischke, M.S., Mahonen, A.P., Miyawaki, K., Hashimoto, Y., Seki, M., Kobayashi, M., Shinozaki, K., Kato, T., Tabata, S., et al. (2004). In planta functions of the *Arabidopsis* cytokinin receptor family. Proc. Natl. Acad. Sci. USA *101*, 8821–8826.

Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics *17*, 754–755.

Huot, B., Yao, J., Montgomery, B.L., and He, S.Y. (2014). Growth-defense tradeoffs in plants: a balancing act to optimize fitness. Mol. Plant *7*, 1267–1287.

Hyatt, D., Chen, G.-L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics *11*, 119.

Jones, J.D., and Dangl, J.L. (2006). The plant immune system. Nature *444*, 323–329.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. *44* (D1), D457–D462.

Kawaharada, Y., Kelly, S., Nielsen, M.W., Hjuler, C.T., Gysel, K., Muszyński, A., Carlson, R.W., Thygesen, M.B., Sandal, N., Asmussen, M.H., et al. (2015). Receptor-mediated exopolysaccharide perception controls bacterial infection. Nature *523*, 308–312.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Kunze, G., Zipfel, C., Robatzek, S., Niehaus, K., Boller, T., and Felix, G. (2004). The N terminus of bacterial elongation factor Tu elicits innate immunity in *Arabidopsis* plants. Plant Cell *16*, 3496–3507.

Lagesen, K., Hallin, P., Rødland, E.A., Staerfeldt, H.-H., Rognes, T., and Ussery, D.W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. *35*, 3100–3108.

Lakshmanan, V., Castaneda, R., Rudrappa, T., and Bais, H.P. (2013). Root transcriptome analysis of *Arabidopsis thaliana* exposed to beneficial *Bacillus subtilis* FB17 rhizobacteria revealed genes for bacterial recruitment and plant defense independent of malate efflux. Planta *238*, 657–668.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Li, L., Li, M., Yu, L., Zhou, Z., Liang, X., Liu, Z., Cai, G., Gao, L., Zhang, X., Wang, Y., et al. (2014). The FLS2-associated kinase BIK1 directly phosphorylates the NADPH oxidase RbohD to control plant immunity. Cell Host Microbe *15*, 329–338.

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., Li, Y., Li, S., Shan, G., Kristiansen, K., et al. (2010). De novo assembly of human genomes with massively parallel short read sequencing. Genome Res. *20*, 265–272.

Liang, Y., Cao, Y., Tanaka, K., Thibivilliers, S., Wan, J., Choi, J., Kang, C.h., Qiu, J., and Stacey, G. (2013). Nonlegumes respond to rhizobial Nod factors by suppressing the innate immune response. Science *341*, 1384–1387.

Lozano-Durán, R., and Zipfel, C. (2015). Trade-off between growth and immunity: role of brassinosteroids. Trends Plant Sci. *20*, 12–19.

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

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.

Moulin, L., Munive, A., Dreyfus, B., and Boivin-Masson, C. (2001). Nodulation of legumes by members of the beta-subclass of Proteobacteria. Nature *411*, 948–950.

Okazaki, S., Kaneko, T., Sato, S., and Saeki, K. (2013). Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. Proc. Natl. Acad. Sci. USA *110*, 17131–17136.

Okushima, Y., Overvoorde, P.J., Arima, K., Alonso, J.M., Chan, A., Chang, C., Ecker, J.R., Hughes, B., Lui, A., Nguyen, D., et al. (2005). Functional genomic analysis of the auxin response factor gene family members in *Arabidopsis thaliana*: unique and overlapping functions of ARF7 and ARF19. Plant Cell *17*, 444–463.

Oldroyd, G.E.D. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat. Rev. Microbiol. *11*, 252–263.

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

Pagel, M. (1994). Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. Proc. Biol. Sci. *255*, 37–45.

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. Bioinformatics *20*, 289–290.

Perilli, S., and Sabatini, S. (2010). Analysis of root meristem size development. Methods Mol. Biol. *655*, 177–187.

Poitout, A., Martiniere, A., Kucharczyk, B., Queruel, N., Silva-Andia, J., Mashkoor, S., Gamet, L., Varoquaux, F., Paris, N., Sentenac, H., et al. (2017). Local signalling pathways regulate the *Arabidopsis* root developmental response to *Mesorhizobium loti* inoculation. J. Exp. Bot. *68*, 1199–1211.

Poupin, M.J., Greve, M., Carmona, V., and Pinedo, I. (2016). A complex molecular interplay of auxin and ethylene signaling pathways is involved in *Arabidopsis* growth promotion by *Burkholderia phytofirmans* PsJN. Front. Plant Sci. *7*, 492.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS One *5*, e9490.

Pupko, T., Pe'er, I., Shamir, R., and Graur, D. (2000). A fast algorithm for joint reconstruction of ancestral amino acid sequences. Mol. Biol. Evol. *17*, 890–896.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

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

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

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Ronquist, F., and Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics *19*, 1572–1574.

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.

Schlaeppi, K., Dombrowski, N., Garrido-Oter, R., Ver Loren van Themaat, E., and Schulze-Lefert, P. (2014). Quantitative divergence of the bacterial root microbiota in *Arabidopsis thaliana* relatives. Proc. Natl. Acad. Sci. USA *111*, 585–592.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*, 539.

Stringlis, I.A., Proietti, S., Hickman, R., Van Verk, M.C., Zamioudis, C., and Pieterse, C.M.J. (2018). Root transcriptional dynamics induced by beneficial rhizobacteria and microbial immune elicitors reveal signatures of adaptation to mutualists. Plant J. *93*, 166–180.

Tritt, A., Eisen, J.A., Facciotti, M.T., and Darling, A.E. (2012). An integrated pipeline for de novo assembly of microbial genomes. PLoS One *7*, e42304.

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

van Velzen, R., Holmer, R., Bu, F., Rutten, L., van Zeijl, A., Liu, W., Santuari, L., Cao, Q., Sharma, T., Shen, D., et al. (2018). Comparative genomics of the nonlegume *Parasponia* reveals insights into evolution of nitrogen-fixing rhizobium symbioses. Proc. Natl. Acad. Sci. USA *115*, E4700–E4709.

Wu, M., and Eisen, J.A. (2008). A simple, fast, and accurate method of phylogenomic inference. Genome Biol. *9*, R151.

Yang, Z. (2006). Computational Molecular Evolution (Oxford University Press Oxford).

Yeoh, Y.K., Dennis, P.G., Paungfoo-Lonhienne, C., Weber, L., Brackin, R., Ragan, M.A., Schmidt, S., and Hugenholtz, P. (2017). Evolutionary conservation of a core root microbiome across plant phyla along a tropical soil chronosequence. Nat. Commun. *8*, 215.

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. (2006). The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. Genome Biol. *7*, R34.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287.

Zamioudis, C., Mastranesti, P., Dhonukshe, P., Blilou, I., and Pieterse, C.M.J. (2013). Unraveling root developmental programs initiated by beneficial *Pseudomonas* bacteria. Plant Physiol. *162*, 304–318.

Zgadzaj, R., Garrido-Oter, R., Jensen, D.B., Koprivova, A., Schulze-Lefert, P., and Radutoiu, S. (2016). Root nodule symbiosis in *Lotus japonicus* drives the establishment of distinctive rhizosphere, root, and nodule bacterial communities. Proc. Natl. Acad. Sci. USA *113*, E7996–E8005.

Zhao, C.Z., Huang, J., Gyaneshwar, P., and Zhao, D. (2017). Rhizobium sp. IRBG74 alters *Arabidopsis* root development by affecting auxin signaling. Front. Microbiol. *8*, 2556.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
Plant-associated bacterial strains	See Table S1	See Table S1
Chemicals, Peptides, and Recombinant Proteins		
flg22 (Flagelin 22)	EZBiolab	Cat# cp7201
Streptomycin sulfate salt,powder	Sigma-Aldrich	Cat# S6501
Tryptone/Peptone ex casein	Carl Roth	Cat# 8952
Yeast extract	BD Biosciences	Cat# 212750
Agar, granulated	BD Biosciences	Cat# 214530
Tryptic Soy Broth	Sigma-Aldrich	Cat# T8907
X-Gluc (5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid, cyclohexylammonium salt)	Thermo Scientific	Cat# R0852
Deposited Data		
Bacterial genome sequences	https://www.ebi.ac.uk/ena/	PRJEB26998
Raw RNA-seq data	https://www.ebi.ac.uk/ena/	PRJEB27007
Experimental Models: Organisms/Strains		
Arabidopsis thaliana wild-type, Columbia (Col-0)	TAIR	CS60000
A. thaliana wild-type Bayreuth (Bay-0)	TAIR	CS6608
A. thaliana wild-type St. Maria d. Feiria (Fei-0)	TAIR	CS22645
A. thaliana wild-type Drahonin (Dra-2)	TAIR	CS1120
A. thaliana Wave_131Y	TAIR	CS781665
A. thaliana ahk3-3	Higuchi et al., 2004	N/A
A. thaliana arr1-3 arr12-1	TAIR	CS6981
A. thaliana tir1-1	TAIR	CS3798
A. thaliana aux1-7	TAIR	CS3074
A. thaliana axr4-2	TAIR	CS8019
A. thaliana arf7-1 arf19-1	Okushima et al., 2005	N/A
A. thaliana fls2	TAIR	CS875982
A. thaliana proCYP71A12::GUS	Millet et al., 2010	N/A
Oligonucleotides		
Primers used in this study	See Table S6	See Table S6
Software and Algorithms		
Trimmomatic	Bolger et al., 2014	http://www.usadellab.org/cms/?page=trimmomatic
A5	Tritt et al., 2012	https://sourceforge.net/p/ngopt/wiki/A5PipelineREADME/
SOAPdenovo	Li et al., 2010	http://soap.genomics.org.cn/soapdenovo.html
HGAP	Chin et al., 2013	https://github.com/PacificBiosciences/Bioinformatics-Training/wiki/HGAP
HMMER	Eddy, 2011	http://hmmer.org
OrthoFinder	Emms and Kelly, 2015	https://github.com/davidemms/OrthoFinder
USEARCH	Edgar, 2013	https://www.drive5.com/usearch/
QIIME	Caporaso et al., 2010	http://qiime.com
UCHIME	Edgar et al., 2011	https://www.drive5.com/uchime/
CLUSTALO	Sievers et al., 2011	https://www.ebi.ac.uk/Tools/msa/clustalo/
MrBayes	Ronquist and Huelsenbeck, 2003	http://mrbayes.sourceforge.net/
FastTree	Price et al., 2010	http://www.microbesonline.org/fasttree/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
FASTX-Toolkit	N/A	http://hannonlab.cshl.edu/fastx_toolkit/
TopHat2	Kim et al., 2013	http://ccb.jhu.edu/software/tophat/index.shtml
Bowtie2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
BEDtools	Quinlan and Hall, 2010	http://bedtools.readthedocs.io/en/latest/index.html
ImageJ	Schneider et al., 2012	https://imagej.nih.gov/ij/
R	N/A	https://www.r-project.org/
Plotting and statistical analyses (custom scripts)	N/A	http://www.mpipz.mpg.de/R_scripts

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Paul Schulze-Lefert (schlef@mpipz.mpg.de).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacterial Strains

The bacterial strains used in this study are summarized in Tables S1 and S4. NGR234 and R7A, CIAT899, and USDA61 were kindly provided by S. Radutoiu (Aarhus University, Denmark), R. Geurts (Wageningen University, Netherlands), and S. Okazaki (Tokyo University of Agriculture and Technology, Japan), respectively.

Plant Model

A. thaliana Col-0 wild-type (N60000) and transgenic line Wave_131Y (N781665) were obtained from the Nottingham Arabidopsis Stock Centre (NASC). Natural accessions of *A. thaliana* (Bay-0, Fei-0, and Dra-2), cytokinin receptor (*ahk3-3*) and signaling mutants (*arr1-3arr12-1*), auxin receptor (*tir1-1*) and transporter (*aux1-7* and *axr4-2*) mutants, auxin signaling mutant (*arf7-1arf19-1*), *fls2* mutant (SAIL_691C4), and a transgenic line expressing a proCYP71A12::GUS construct were kindly provided by K. Schlaeppi (Agroscope, Switzerland), S. Sabatini (Sapienza University of Rome, Italy), E. Kombrink (MPIPZ, Germany), M. Bennett (University of Nottingham, UK), K. Tsuda (MPIPZ, Germany), and F. Ausubel (The Massachusetts General Hospital, USA), respectively.

Growing Conditions for Plant Models

Seeds were surface-sterilized in 70% ethanol for 10 min followed by a brief wash with 100% ethanol. Plants were grown on agar plates prepared with half-strength Murashige and Skoog (MS) medium (750 μM $MgSO_4$, 625 μM KH_2PO_4, 10.3 mM NH_4NO_3, 9.4 mM KNO_3, 1.5 mM $CaCl_2$, 55 nM $CoCl_2$, 53 nM $CuCl_2$, 50 μM H_3BO_3, 2.5 μM KI, 50 μM $MnCl_2$, 520 nM Na_2MoO_4, 15 μM $ZnCl_2$, 75 μM Fe-EDTA, 500 μM MES-KOH pH 5.5) supplemented with 1% Difco Agar Granulated (BD Biosciences, USA), an agar which contains limited amounts of bioavailable nutrients (Gruber et al., 2013). Phosphate- and nitrogen-depleted media were prepared according to a previous study (Gruber et al., 2013). Plants were grown under short-day conditions at 21°C under light (10 hr) and at 19°C under dark (14 hr).

Culture Conditions for *In Vitro* Systems

Bacterial strains were routinely cultured in TY (5 g/L tryptone, 3 g/L yeast extract, 10 mM $CaCl_2$) or 50% TSB media (Sigma-Aldrich, USA) and stored in 16% glycerol at −80°C.

METHOD DETAILS

Isolation and Sequencing of Bacterial Strains

Surface-sterilized *A. thaliana* seeds were sown at a density of four plants per pot (7×7×9 cm) and stratified for 3-4 days. Plants were grown in the greenhouse for 6 weeks under short day conditions (8/16 hr day/night with temperatures of 22°C/18°C and relative humidity of 70%). After harvest, roots were mechanically separated from the adhering soil particles and a defined root segment of 3 cm, beginning 0.5 cm distal from the hypocotyl, was sampled. Soil particles still attached to the roots were removed by gentle tapping. Roots were collected in 15-mL conical tubes containing 5 mL phosphate-buffered saline (PBS)-S buffer (130 mM NaCl, 7 mM Na_2HPO_4, 3 mM NaH_2PO_4, pH 7.0, 0.02% Silwet L-77) and washed for 15 min at 180 rpm on a shaker. Roots were transferred to a new 15-mL conical tube and the remaining soil particles were centrifuged for 20 min at 15,000 × *g*. After washing for a second time, roots were transferred to a new 15-mL conical tube and sonicated with ten cycles of a 30 s pulse at 160 W followed by a 30 s interval (Bioruptor Next Gen UCD-300; Diagenode, Belgium). Roots were transferred to a 1.5-mL tube containing 10 mM $MgCl_2$ and

mechanically disrupted using sterile 3-mm metal beads and a Precellys24 tissue homogenizer (Bertin, France) by three cycles of 5,000 rpm for 30 s. The supernatant was collected by centrifugation at 1,000 × *g* for 5 min and serial dilutions were plated onto flour and TWYE media (flour: 6 g/L flour, 0.3 g/L yeast extract, 0.3 g/L sucrose, 0.3 g/L $CaCO_3$, 1.8% agar; TWYE: 0.25 g/L yeast extract, 0.5 g/L K_2HPO_4, 1.8% agar), including 50 μg/mL benzimidazole to inhibit fungal growth. Plates were incubated for 3-4 days at 28°C. Emerged single colonies were transferred with a sterile pipette tip to 400 μL liquid media in a 96-well format and incubated for up to 7 days at 28°C at 200 rpm. For taxonomic identification, 100 μL of the culture was transferred to a 96-well PCR plate, bacterial cells were heat-disrupted for ten min at 100°C, and cell debris was removed by centrifugation at 3,000 rpm for 10 min. The supernatant was used for PCR amplification of 16S rRNA V5-V8 regions using the primers 799F (5′-AACMGGATTAGATACCCKG-3′) and 1392R (5′-ACGGGCGGTGTGTRC-3′). PCR reactions were performed under sterile conditions using 3 μL of the supernatant in a total volume of 25 μL that contained 1.25 U DFS-Taq DNA Polymerase (Bioron, Germany), 1 × complete reaction buffer, 0.3% BSA, 200 μM of each dNTP, and 400 nM of each primer, using the following PCR program: 94°C for 5 min, 35 cycles of 94°C for 30 s, 50°C for 30 s and 72°C for 30 s, followed by 72°C for 5 min. PCR quality was assessed by visualizing 5 μL of the amplicon on a 1% TAE-agarose gel with ethidium bromide. PCR products were purified using the QIAquick PCR Purification Kit (Qiagen, Germany), DNA concentrations were measured using a Nanodrop spectrophotometer (PeqLab, Germany) and adjusted to 20 ng/μL, and the purified products were subjected to Sanger sequencing. Genomic DNA of each strain was extracted and purified as described previously (Bai et al., 2015) and sequenced using a combination of the Illumina HiSeq 2500 (Illumina, USA) and PacBio RS II (Pacific Biosciences, USA) platforms (see below). All sequencing was performed at the Max Planck Genome Center (Cologne, Germany) or at AgBiome (Research Triangle Park, USA).

Bacterial Inoculation and Image Analysis

Bacteria strains were cultured for 2-3 days until strains reached the stationary phase in their respective liquid media (see Table S4) at 28°C (180-200 rpm). Bacterial cultures in the stationary phase were diluted 5-fold in sterile liquid media and pre-cultured for another 2-4 hr in order to metabolically activate bacterial cells. Bacterial cells were collected and resuspended in 10 mM $MgCl_2$ and adjusted to an optical density of 0.5 at 600 nm (OD_{600nm}). The bacterial resuspension mixture was added to plant MS media cooled to 40°C -45°C to obtain a final OD_{600nm} of 0.00005, which corresponds to approximately 5×10^4 cells/mL. For the root growth assay of a panel of Alphaproteobacteria (Figure 3) and the developmental analysis with *Rhizobium* 129E (Figures 4 and S6), surface sterilized seeds were directly planted onto the bacteria-containing medium and grown for 5-27 days. For the analyses using three *A. thaliana* accessions (Figures S5A and S5B) and phosphorus- and nitrogen-deficient conditions (Figures S5C–S5E), as well as for the RNA-seq experiment (Figures 5 and S12), plants were precultured on half-strength MS medium supplemented with 1% sucrose for 6 days before seedlings were transferred onto bacteria-containing media under sterile conditions and grown for 18 days. For flg22 treatment (Figures 6 and 7), the mock or bacteria-containing medium was supplemented with 1 μM flg22 peptide (EZBiolab, USA). Wild-type Col-0 or the *fls2* mutant were pre-grown on half-strength MS medium supplemented with 1% sucrose and 1% agar for 5-7 days and transferred to bacteria-containing medium with sterile toothpicks. Plants were grown for 14-15 days after transfer. For co-inoculation assays, each bacterial strain was added in an equal volume (1/10,000 volume at OD_{600}=0.5), resulting in a doubled amount of total bacterial titer compared to the mono-association assays.

DNA Extraction and Quantitative PCR

Roots inoculated with bacteria were collected from plates and briefly washed two times in sterile water. After removal of excess water with a glass fiber filter paper (Whatman, UK), roots were frozen in liquid nitrogen and stored at -80°C until further processing. DNA was extracted using Lysing Matrix E (MP Biomedicals, USA) and the DNeasy Plant Mini Kit (Qiagen, Germany) according to the manufacturer's instructions. Quantitative PCR was performed in a 15 μL reaction mixture containing 7.5 μL iQ SYBR Green Supermix (BIO-RAD, USA), 0.8 μM forward primer, 0.8 μM reverse primer, and 2-4 μL DNA template. PCR was performed using a CFX Connect (Bio-Rad, USA) using the following cycles: 95°C for 3 min, followed by 40 cycles of 95°C for 10 s, 60°C for 30 s, and 72°C for 30 s. The delta-Ct method was used to estimate the relative abundance of bacteria to the abundance of plant DNA. Primers used in this study are shown in Table S6.

Confocal Microscopy and Image Analysis

Transgenic *A. thaliana* line Wave_131Y, expressing NPSN12 fused with the enhanced yellow fluorescent protein (EYFP), was observed under a CLSM780 confocal microscope (Zeiss, Germany) equipped with a water-immersion 20× objective using the 514-nm line of a 25 mW argon ion laser. All image analysis was performed using ImageJ (Schneider et al., 2012). Plates were scanned and marked prior to the microscopy, allowing correlations between the parameters of individual plants.

RNA Extraction and RNA-seq

For transcriptomic analysis, 6-day-old *A. thaliana* seedlings were inoculated with *Rhizobium* 129E or mock control under phosphate-sufficient or phosphate-depleted conditions and incubated for 4, 8, 12, and 16 days under short-day conditions, as described above. Roots from three plates were combined for one replicate and a total of three replicates were sampled for each condition. Roots were homogenized with Lysing Matrix E and TissueLyser II (20 beats per second for 30 s; Qiagen, Netherlands) and RNA was extracted with the miRNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. RNA quality was determined using a 2100 Bioanalyzer (Agilent Technologies, USA). Preparation of Illumina sequencing libraries was conducted by the Max Planck Genome

Center) using an input of 1.5 μg of total RNA. Sequences were generated using the Illumina HiSeq2500 platform. Approximately 20M reads per sample with a length of 100 bp were generated and further analyzed.

GUS Histochemical Assay

GUS staining was performed as described previously (Millet et al., 2010) with slight modifications. Briefly, approximately 15 *A. thaliana* seeds were germinated in a six-well culture plate in half-strength MS liquid medium (0.5% sucrose) and incubated for 7 days under short-day conditions at 21°C under light (10 hr) and at 19°C under dark (14 hr). Afterwards, the growth medium was replaced by fresh liquid MS medium without sucrose. Bacterial cells actively growing in 50% TSA medium were collected, washed once, and then resuspended with 10 mM $MgSO_4$. OD_{600nm} was measured by a spectrometer. Bacterial suspensions (final OD_{600nm}=0.1), flg22 (final concentration of 100 nM), or an equal volume of 10 mM $MgSO_4$ was added to the liquid medium in the six-well plates and incubated for 24 hr under the same conditions. Treated seedlings were washed with 50 mM sodium phosphate buffer (pH 7.0) and incubated with GUS substrate solution (50 mM sodium phosphate buffer [pH 7.0], 10 mM EDTA, 0.5 mM $K_4[Fe(CN)_6]$, 0.5 mM $K_3[Fe(CN)_6]$, 0.5 mM X-Gluc, and 0.01% Silwet L-77) at 37°C for 2-3 hr after 5 min of vacuum infiltration. Tissues were fixed twice with a 3:1 ethanol:acetic acid solution and cleared in lactic acid. Images were taken using an Axio Imager A2 microscope (Zeiss).

QUANTIFICATION AND STATISTICAL ANALYSIS

Genome Assembly

Genomes sequenced using the Illumina platform with an insert size of 350 bp and an approximate depth of 5M reads per strain were quality-filtered by passing short reads through a quality and length-trimming filter using Trimmomatic (Bolger et al., 2014) with the default parameters and subsequently assembling them using an ensemble of two approaches: A5 (Tritt et al., 2012) and SOAPdenovo (Li et al., 2010). In each case, the assembly with the smaller number of scaffolds was selected. Contigs smaller than 1,000 bases were removed from the assemblies. Additionally, a subset of genomes were sequenced using the Pacific Biosciences platform and these long reads were assembled into complete genomes using the Hierarchical Genome Assembly Process (HGAP) assembler (Chin et al., 2013). Detailed statistics concerning the assemblies of all genomes analyzed in this study can be found in Table S1. Accession numbers corresponding to the raw reads of all genomes, including those of previously published organisms, can likewise be found in Table S1.

Genome Annotation and Orthology Inference

To determine homology relationships between gene sequences, we first predicted putative protein-coding sequences using Prodigal (Hyatt et al., 2010). Annotation of candidate Open Reading Frames (ORFs) was then conducted using the KEGG Orthology (KO) database (Kanehisa et al., 2016) as previously described (Bai et al., 2015). Briefly, sequences from each KO were aligned using Clustal Omega (Sievers et al., 2011) and hidden Markov models (HMMs) generated based on each multiple sequence alignment using the HMMER suite (Eddy, 2011). Subsequently, the HMMs were employed to search all putative protein-coding sequences. Sequences matching a given KO group with an E value lower than 10×10^{-5} with a coverage of at least 70% of the total length were assigned to that KO group. All sequences assigned to the same KO were considered homologous and classified as belonging to the same Cluster of Orthologous Groups (COG). For the pangenome analysis and permutation-based analysis of saturation curves we also conducted a more comprehensive prediction of COGs using OrthoFinder (Emms and Kelly, 2015), a sequence-based *de novo* method for orthogroup inference. Subsequently, for each COG we generated a phyletic pattern consisting of a binary vector of presence/absence for each KO group or COG in each genome of the dataset.

Natural Community Amplicon Sequencing Meta-Analysis

First, we retrieved from the public databases all raw, unprocessed sequences corresponding to the analyzed 16S rRNA gene surveys (Bai et al., 2015; Bulgarelli et al., 2015; Schlaeppi et al., 2014; Zgadzaj et al., 2016) and compiled a master meta-data table containing information of all samples across studies (Table S2). Next, a database of high-quality reference sequences was built by extracting 16S rRNA gene sequences from the Rhizobiales genome assemblies using RNAmmer (Lagesen et al., 2007). Identical reference sequences were dereplicated prior to downstream analyses using usearch. Subsequently, 16S rRNA gene sequences were processed using a combination of custom scripts as well as tools from the QIIME (Caporaso et al., 2010) and USEARCH (Edgar, 2013) pipelines. First, reads were truncated to an even length (290 bp) using the truncate *fasta_qual_files.py* QIIME script. Libraries were demultiplexed (*split_libraries.py*) and only reads with a quality score above 25 were retained for subsequent analysis. After dereplication and removal of singletons we conducted a reference-based clustering of sequences into operational taxonomic units (OTUs) using the UPARSE algorithm (Edgar, 2013) at 99% identity using high-quality 16S dereplicated sequences extracted from the whole-genome assemblies. In parallel, we conducted *de novo* clustering of all sequences into OTUs using the standard UPARSE algorithm at 97% identity. After chimeric sequences were filtered using UCHIME (Edgar et al., 2011), we used all non-artifactual OTUs per sample to normalize Rhizobiales abundances as a percentage of the total sample size (Figure 1A). The resulting OTU table was used in all subsequent statistical analyses of differentially abundant taxa as well analyses of diversity (Figure 1B).

Comparative Genomics and Ancestral Character Reconstruction

Inference of an accurate species tree was conducted as follows: first, from each genome, we extracted a set of up to 32 single-copy phylogenetic marker genes that are present in the majority of sequenced bacterial genomes called AMPHORA genes (Wu and Eisen, 2008) using custom-built HMMs and the hmmsearch tool (Eddy, 2011). The resulting sequences were concatenated independently for each genome and then aligned using Clustal Omega (Sievers et al., 2011). For each dataset, rooted species trees were generated from the AMPHORA multiple sequence alignments using MrBayes (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) with a strict molecular clock and a general time reversible model (GTR+G+I; 10,800,000 iterations, parameters *nst=6, rates=invgamma* and *brlenspr=clock:uniform*). Similar results as reported above were obtained by reconstructing the species tree using an alternative maximum likelihood approach that is implemented in FastTree (Price et al., 2010) with a general time reversible model of DNA evolution and without a molecular clock constraint (data not shown). Polytomies in the trees were resolved by inserting branches of zero length to avoid conflicts in the inference of ancestral characters. Given the matrix of phyletic patterns and the estimated species tree, we used a maximum likelihood approach (Pagel, 1994) for the estimation of ancestral states using the implementation provided with the ape package (Paradis et al., 2004) for R software (https://www.r-project.org/), which employs a 'two-pass' algorithm for the joint estimation of the likelihood of the ancestral states (Yang, 2006). This approach gives similar results as stochastic mapping (Paradis et al., 2004; Pupko et al., 2000) while being much faster, thus enabling the estimation of ancestral characters for very large datasets (several thousands of genomes or more).

Analyses of functional diversity between sequenced isolates were conducted as previously described (Bai et al., 2015). First, we generated a presence/absence profile of each KO group (or phyletic pattern) for each genome in the data set. Subsequently, a distance measure based on the Pearson correlation of each pair of phyletic patterns was calculated which allowed us to embed each genome as a data point in a metric space.

Image Analysis for Root Growth Quantification

Plate images were taken and root length was measured using ImageJ (Schneider et al., 2012). Shoot fresh weights were measured after removal of excess water by a paper towel. Each biological replicate contained at least three technical replicates. Primary root lengths and shoot fresh weights were normalized to the mock control within each replicate. Statistical tests were performed using R software according to Dunnett's or Tukey's HSD tests depending on the dataset (see figure legends).

Image Analysis for Root Developmental Analysis

Transitions from the meristematic zone (MZ) to the elongation zone (EZ) was determined as described previously (Perilli and Sabatini, 2010). The number of cortex cells within the MZ was counted until the area where cells began to elongate after transition (see arrowheads in Figure 4A). Transition from the EZ to the differentiation zone (DZ) was determined by the first appearance of a root hair in the epidermal cell layer, and the longitudinal length of the cortex cell underneath was measured accordingly (see dotted lines in Figures S6A and S6B). For each root, two Z stack images were randomly taken from the DZ and the longitudinal length of all cortex cells was measured (n = 638, average of 23.6 cells/root), of which mean values were used as the cell length in the DZ.

RNA-seq Data Analysis

Obtained sequences were quality filtered by trimming the Illumina adaptor and discarding every sequence with a quality score below 25 using the FASTX toolkit. High quality sequences were mapped to the *A. thaliana* genome (TAIR10 annotation file) TopHat2 (Kim et al., 2013) and Bowtie2 (Langmead and Salzberg, 2012). A count table for each gene and sample was then generated using the BEDtools suite (function coverageBED) (Quinlan and Hall, 2010).

All statistical analyses were performed using R. Read count data from Castrillo et al. (2017) were retrieved from the Gene Expression Omnibus (GEO: GSE87336). Genes were removed from the analysis unless at least 100 counts over all samples and more than ten counts at least in two samples were detected. Normalization using the library size-dependent trimmed mean of M-values (TMM) method, calculation of $\log_2$ counts per million and fold changes, and statistical analyses of differentially expressed genes were completed using the edgeR R package (Robinson et al., 2010). Median-centered Z scores based on $\log_2$ counts per million were used for *k*-means clustering, for which *k* was objectively determined by computing Bayesian information criterion. Differentially expressed genes were identified by fitting a negative binomial generalized linear model to the genes using the *glmFit* function with the following formula: $\sim$ 0 + treatment*nutrient*timepoint + replicate. P values were calculated using the *glmLRT* function and corrected for multiple tests controlling for the false discovery rate (FDR) with $\alpha = 0.05$ using the *decideTestsDGE* function. Cluster-wise gene ontology enrichment analysis was performed with the clusterProfiler R package using *compareCluster* and *enrichGO* functions (Yu et al., 2012).

DATA AND SOFTWARE AVAILABILITY

Whole-genome assemblies and transcriptome sequencing reads have been deposited in the European Nucleotide Archive (ENA) under the accession numbers ENA: PRJEB26998, PRJEB27007, respectively. To ensure reproducibility, the scripts used for computational analyses and the corresponding raw and intermediate data will be made available at http://www.mpipz.mpg.de/R_scripts.