Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly

Kateryna Zhalnina[1,2], Katherine B. Louie[1], Zhao Hao[2], Nasim Mansoori[1,3], Ulisses Nunes da Rocha[2,4], Shengjing Shi[5], Heejung Cho[2,6], Ulas Karaoz[2], Dominique Loqué[1,3,6,7], Benjamin P. Bowen[1], Mary K. Firestone[2,8], Trent R. Northen[1*], and Eoin L. Brodie[2,8*]

Affiliations

[1]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA.

[2]Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, California, USA.

[3]Joint BioEnergy Institute, Biosystems Engineering Division, Lawrence Berkeley National Laboratory, Emeryville, California.

[4]Department of Environmental Microbiology, Helmholtz Centre for Environmental Research - UFZ, Leipzig, Germany.

[5]Lincoln Science Centre, AgResearch Ltd, Christchurch, New Zealand.

[6]Department of Plant and Microbial Biology, University of California, Berkeley, California, USA.

[7]Université Claude Bernard Lyon 1, INSA de Lyon, CNRS, UMR5240, Microbiologie, Adaptation et Pathogénie, Villeurbanne, France

[8]Department of Environmental Science, Policy and Management, University of California, Berkeley, California, USA.

([*]Corresponding authors: trnorthen@lbl.gov, elbrodie@lbl.gov)

Abstract

Like all higher organisms, plants have evolved in the context of a microbial world, shaping both their evolution and their contemporary ecology. Interactions between plant roots and soil microorganisms are critical for plant fitness in natural environments. Given this co-evolution and the pivotal importance of plant-microbial interactions, it has been hypothesized, and a growing body of literature suggests, that plants may regulate the composition of their rhizosphere to promote the growth of microorganisms that improve plant fitness in a given ecosystem. Here, using a combination of comparative genomics and exometabolomics, we show that pre-programmed developmental processes in plants (*Avena barbata*) result in consistent patterns in the chemical composition of root exudates. This chemical succession in the rhizosphere interacts with microbial metabolite substrate preferences that are predictable

from genome sequences. Specifically, we observed a preference by rhizosphere bacteria for consumption of aromatic organic acids exuded by plants (nicotinic, shikimic, salicylic, cinnamic, and indole-3-acetic). The combination of these plant exudation traits and microbial substrate uptake traits interact to yield the patterns of microbial community assembly observed in the rhizosphere of an annual grass. This discovery provides a mechanistic underpinning for the process of rhizosphere microbial community assembly and provides an attractive direction for the manipulation of the rhizosphere microbiome for beneficial outcomes.

Introduction

The area surrounding growing plant roots in soil (the rhizosphere) represents a critical hotspot for biogeochemical transformation that underlies the process of soil formation, carbon cycling and the ultimate productivity of Earth's terrestrial ecosystems. Within the rhizosphere, complex and dynamic interactions between plants and networks of organisms, particularly microorganisms have been shaped by over 450 million years of co-evolution. With such evolutionary optimization, it is not surprising that somewhat consistent patterns have emerged. The 'rhizosphere effect'[1] describes the enrichment of microbial cells and activity near growing roots and has been shown to involve the "selection" of phylogenetically related microorganisms, with plants of different species, geographic locations, climates, and land management showing distinct rhizosphere microbiomes[2-9]. Across systems, some consistent trends have also been observed, for example an increase in bacteria of the Alphaproteobacteria sub-phylum was reported in rhizosphere soil of a variety of plants[2,10-12] . Conversely, the abundance of Actinobacteria has been shown to decrease during late developmental stages[2,3,12]. Several specific traits have been identified and associated with rhizosphere-enrichment, such as the presence of secretory systems, adhesion, phage defense, iron mobilization and sugar transport[11,13]. Clearly the rhizosphere microbial community structure is the result of a complex series of interactions and feedbacks between plant roots, microorganisms and the soil physical and chemical environment. However, despite the growing number of studies demonstrating that plant development influences the composition of soil microbiome and its functional capacity, relatively few studies[14-16] have sought to understand the molecular and chemical basis of the role of dynamic plant exudation in the establishment of rhizosphere microbiota.

Plants exude a variable but substantial amount (11-40%) of photosynthesis-derived carbon (C) creating a diverse chemical milieu[17,18]. Exuded compounds include sugars, amino acids, organic acids, fatty acids and secondary metabolites[18-20]. The composition of root exudates is not a uniform nor static property and varies depending on plant species, developmental stage, root traits, environmental conditions, nutrition, soil type, etc.[18,21-24] Released compounds have been shown to attract beneficial microorganisms and influence the assembly of rhizosphere microbiomes that enhance the capacity of plants to adapt to their environment[25].

Many studies have demonstrated the impact of small signaling molecules (acyl-homoserine lactones, flavonoids, non-proteinogenic amino acids, etc.)[8,26-28], polymers[29], antimicrobials[28,30] or plant hormones, such as salicylic acid[16] on the interactions between plants and microorganisms in the rhizosphere; however these compounds represent only a

small fraction of exuded metabolites. It is not clear how the interaction between root exudate chemistry and microbial substrate preferences combine to influence rhizosphere community assembly and succession. Are there relationships between the dynamics of exudate composition and the growth of specific soil microorganisms? If so, can those relationships be predicted and generalized?

To address this uncertainty we integrate information from comparative genomics and a recently described exometabolomics approach[31,32] to explore the metabolic potential of soil bacteria, the composition of root exudates produced by an annual grass through its developmental stages, and the substrate uptake preferences of isolated soil bacteria representing groups that display distinct successional responses to growing plant roots. We hypothesized that bacteria enriched in the rhizosphere have distinct substrate preferences relative to those bacteria that are not enriched or decline in response to growing plant roots. We demonstrate that growth responses of bacteria in the rhizosphere can be explained by their predicted and observed substrate preferences and the chemical composition of root exudates, thus providing evidence of direct manipulation of the soil microbiome through the specific composition of exudates.

Results and discussion

Succession of bacterial isolates in the rhizosphere during growth of *Avena*. The "rhizosphere effect" has been observed across countless plant species and soil types[2,4,10,11,33]. The bulk soil represents a seed bank of potential organisms that may flourish in response to the resources from a growing root. Numerous studies have shown that selection in the rhizosphere is non-random, with some apparent phylogenetic conservation[2,34]. This suggests that specific inherited traits are being selected by plants, potentially through their chemical modification of the root zone. Here we studied the mechanisms underlying the response of the soil microbiome to *Avena* root growth. Using media with a range of concentrations, nutrients, anti-oxidants, vitamin and co-factor compositions, together with extended incubation times, 289 heterotrophic bacteria were isolated and phylogenetically characterized from the Mediterranean grassland soil in which *Avena* dominates (Fig. 1). These isolates represented seven phyla, coming mostly from the Actinobacteria and Alpha-proteobacteria known to be dominant in this soil[2]. Isolate recovery varied based on media composition (Fig. 1), indicating selection for organisms from distinct niches. Of these isolates, 39 were selected for whole genome sequencing based on their relative abundance in soil and their phylogeny (Supplementary Figures 1, 2 and 3, Supplementary Data 1). Non-redundant OTUs matched to these 39 isolates together represent approximately 10-12% of the total bacterial community in this rhizosphere soil and 17 of these isolates had relative abundance more than 1% of the total bacterial community in this environment (Supplementary Fig. 2 and Supplementary Table 1).

Those bacterial isolates with sequenced genomes were related to 16S rRNA gene OTUs whose successional dynamics had been previously determined in a rhizotron microcosm experiment where rhizosphere soil was sampled at 0, 3, 6, 9, and 12 weeks of *Avena* growth[2] (Supplementary Fig. 2). Based on this analysis the isolates were classified into response groups (Fig. 2). The first group (n=19) contained isolates that increased in relative abundance in response to the plant growth (termed 'positive responders' herein), and the second group (n=8) contained isolates that declined in relative abundance during plant growth

(termed 'negative responders' herein). It should be noted that although changes in relative abundance can be interpreted as changes in relative fitness, they could be due to decreases or increases in other community members, rather than, or in addition to, changes in absolute abundance of the 'responding' organisms. The remainder of the isolates showed no significant change in relative abundance relative to bulk soils of the 12 week period and are termed 'undefined responders'. The positive response group was comprised of isolates related to the Proteobacteria (*Dongia*, *Rhodospirillales*, *Sphingomonas*, *Mesorhizobium*, *Bradyrhizobium, Caulobacter, Burkholderia, Variovorax, Pseudomonas*), a limited number of Actinobacteria (*Mycobacterium* and *Streptomyces*). Isolates from the negative response group were mostly from the Actinobacteria and Firmicutes (*Bacillus*, *Paenibacillus*). In general, the relative abundance of Actinobacteria declined while, Proteobacteria, particularly alpha proteobacteria, increased during plant growth. This observation is consistent with many other successional studies: including *Arabidopsis thaliana*[4], wheat[11], rice[10], switchgrass[33], maize[7], demonstrating some conservation in the restructuring of the rhizosphere microbiome across plant species and soil types.

Functional traits of soil isolates in the context of the life in the rhizosphere. Based on the relative abundance of 16S rRNA gene sequences, the majority of bacteria classified as positive responders could be related to taxa previously known to be associated with the rhizosphere, and in some cases, promote plant growth[35]. This consistent observation of these taxa as being plant-associated across numerous studies again suggests an evolutionary legacy driven by inherited traits favorable to life in the rhizosphere[36]. To identify these traits, we analyzed their genomes for the presence of features that are hypothesized to be important for both rhizosphere growth and soil organic matter transformation (Fig. 3a,b,c,d, Supplementary Fig. 3, Supplementary Data 2). Specifically, we focused on traits associated with the acquisition of carbon substrates, such as macromolecule depolymerization enzymes, monomer transport, in addition to predicted generation times. We analyzed the distribution of these traits across the bacterial rhizosphere response groups.

Conventional wisdom might suggest that the growth strategy of bacteria enriched in the rhizosphere might be adapted towards rapid-growth[17,33], however contrary to our expectations, a majority of bacteria that showed a positive response to plant growth were predicted to have longer generation times based on codon-usage bias, meaning their genomes bear signatures of slower growth rates (Fig. 3a, Supplementary Fig. 3). This prediction was confirmed through laboratory growth rate experiments for the majority of isolates (Supplementary Fig. 4). As slower growing organisms can have higher substrate utilization efficiency[37], perhaps growth efficiency is favored over growth rate in the rhizosphere.

Substrate preference may confer a selective advantage in the rhizosphere and in fact positive and negative responders did differ in their predicted metabolic potential to utilize organic acids (Fig. 3d). Genes coding for organic acid transporters were significantly more abundant in positive responders when corrected for genome size (Fig. 3d). Similarly, the number of amino acid transporters was also skewed towards higher abundance in bacteria with positive response to root growth. Conversely, genes coding for glycoside hydrolases (GH) (primarily β-glucosidases (GH 1, 3, 5), β-xylosidases (GH 43), β-glucanases (GH 16), β-galactosidases (GH 2), glucoamylases (GH 15), α-glucosidases (GH 13) and α-N-acetylgalactosaminidases (GH 9) were more abundant in the genomes of negative responders

(Fig. 3c, Supplementary Data 2) who are presumably better adapted to life outside the living root zone where easily accessible and assimilable substrates (e.g. monomeric molecules) are less available. Together these results suggest that positive and negative responders have features in their genomes that point to differences in the potential of substrate utilization between these two groups and that they occupy distinct niches within the soil. Although genomics can suggest putative metabolic functions, these functions predicted by genome analyses are hypotheses that require experimental confirmation. To determine whether a relationship existed between the genomic potential for uptake of substrate classes found in root exudates and uptake of those substrates from exudate growth medium, we used an exometabolomics approach.

Analysis of *Avena barbata* root exudate metabolite profiles during plant development. Differences in the metabolic potential for the utilization of specific components of exudates suggests that traits related to this may be important to bacterial success in the rhizosphere. The chemical composition of plant exudates therefore may represent a key means of shaping the microbial composition of the rhizosphere. Plants secrete a cocktail of chemicals through their roots that vary in composition over plant developmental stages and nutritional status[38,39], therefore we tested the composition of *A. barbata* exudates throughout its growth stages. For a number of reasons, we chose to use a hydroponic system to analyze the dynamics of *Avena* exudate chemistry. Using hydroponics is a trade-off[38], plants function differently in hydroponics relative to a real soil system, however hydroponics allows precise control over the chemical milieu in which the plant is growing, facilitates collection of sufficient quantities of freshly produced exudates, avoids issues associated with non-uniform mineral sorption of exudate components and can minimize the microbial transformation of exuded metabolites. In this experiment *Avena* plants were grown *in vitro* and exudates were collected during several developmental phases of *Avena* growth, which roughly proceeded along the same time scale as observed in soil. After collection, exudates were analyzed for total organic carbon and a peak in the quantity of organic carbon exuded was observed during the vegetative stage of 9 week-old *Avena* plants (Supplementary Fig. 5). We next assessed the composition of the exudates across developmental stages using liquid chromatography-mass spectrometry (LC-MS). This demonstrated that *A. barbata* exudates were comprised of a broad range of metabolites, including sugars, sugar alcohols, nucleotides, nucleosides, amino-, organic-, fatty acids, plant hormones and compatible solutes (Fig. 4, Supplementary Data 3). Metabolite profiles from the early developmental stage (week 3) were distinct from those at week 6 and 9, and from the late developmental stage corresponding to senescence (week 12) (Fig. 4a). While the exudate metabolite profiles from 6 and 9 week-old *Avena* were similar.

We then determined the metabolites that significantly changed in abundance with *Avena* development (Fig. 4b, Supplementary Fig. 10). At the early developmental stage (week 3), sucrose and homoserine were at greater concentrations relative to other developmental stages (Fig. 4b). Sucrose is the main sugar found in the phloem[40], it can be strongly allocated to the root tip, decreasing with root maturation[41]. At early developmental stages of root growth, sucrose has been noted as an important factor for the development of symbiotic plant-microbe interactions and potential plant-defense mechanisms[42 43].

Weeks 6 and 9 correspond to the vegetative developmental stages of *A. barbata*[2] and displayed the highest overall release of exudates compared to the other stages (Fig. 4b).

Between week 3 and weeks 6 and 9, amino acids and carboxylic acids with aromatic rings showed the greatest increase (Fig. 4b, Supplementary Fig. 10). Plants release a variety of aromatic compounds, as defense mechanisms against pathogens, as signaling molecules and carbon sources for heterotrophs[44]. This increased release of aromatic compounds (vanillic acid, syringic acid, vanillin, ferulic acid) by *Avena fatua* to the rhizosphere soil was reported earlier as a potential mechanism for allelopathy of wild oat roots[23], however it may also be linked to the physiology of rhizosphere microorganisms.

During plant senescence (week 12), a significant increase in the abundance of quaternary ammonium salts (glycine betaine, betonicine, stachydrine) and plant hormones (indole-3-acetic acid, IAA; and abscisic acid, ABA) was observed. Stachydrine (L-proline betaine) and IAA had the most striking change (170 and 40 times fold change, respectively) during *Avena* development (Supplementary Fig. 10). Betaines are widespread in plants and are produced in response to various types of environmental stress to protect to membranes, enzyme activity, and regulate detoxification of reactive oxygen species[45]. ABA has been suggested to play a role in plant metabolism as a stress-response metabolite that promotes senescence-related processes[46]. IAA has been reported to both delay and promote plant senescence[47,48] and an excess of IAA and ABA are known to inhibit root growth[49]. These results indicate that *A. barbata* releases a variety of metabolites that change with root growth and plant developmental stages in a genetically programmed manner and suggest that changes in exudate metabolite composition over time may contribute to the observed successional patterns in the rhizosphere microbiome (Fig. 2).

Although in our study we were able to detect and identify a large number of metabolites present in plant exudates it is still unclear what fraction of the total exudate C is reflected by individual compounds (undetected metabolites and larger water-soluble polymers remain unknown). Therefore, while there may be a significant increase in a particular metabolite, it is important to note that it may only represent a small fraction of the total exudate C pool.

Metabolism of root exudate metabolites by rhizosphere bacterial isolates. To determine whether the substrate preferences of soil bacteria interact with the chemical composition of root exudates to contribute to microbial succession patterns, we selected 16 isolates as representatives of the positive and negative response groups to plant growth (Fig. 2). We determined differences in the substrate preferences of these isolates using an exometabolomics approach[32]. These isolates were cultured in a medium containing pooled exudates, collected across the various *Avena* developmental stages, with uptake of the specific compounds from the medium measured by LC-MS. Metabolite uptake was represented as a percent of a metabolite depleted from the medium by each isolate compared to the control uninoculated medium (Supplementary Data 4). This exometabolomic approach specifically allows substrate preferences to be evaluated as microorganisms are confronted with a choice of substrates within the exudate mix, in contrast with other approaches where substrate utilization is evaluated individually[16,50-53].

Root exudate metabolites were categorized into six chemical classes including amino acids (organic acids containing amino group), nucleotides and nucleosides, sugars, organic acids (organic acids that do not contain an amino group), quaternary amines and fatty acids. Isolates favored during root growth showed significantly higher uptake of amino acids, organic acids, sugars and quaternary amines (Fig. 5). Organic acids and amino acids showed

the most significant differences in uptake between the positive and negative responders. Conversely, isolates with positive response were skewed towards lower uptake of nucleotides and nucleosides compared to negative responders.

These observations of enhanced amino- and organic acid uptake corroborates our finding that rhizosphere enriched bacteria encode a higher number of transporters for organic acids and amino acids in their genomes compared to bacteria that declined in response to root growth (Fig. 3d). This is compelling evidence that selective uptake of organic acids from the mixture of exudates by rhizosphere bacteria interacts with enhanced release of these compounds during vegetative stages of plant growth. This observation of the importance of organic acids for plant-microbial interactions has been noted previously[16,52-54]. For example, addition of organic acids as single substrates significantly improved tomato root colonization by *Pseudomonas*[51].

Further global analysis of root exudate metabolite uptake by isolates (Supplementary Fig. 11) showed that uptake of most root exudate metabolites was similar across isolates with a large percentage of major proteinogenic amino acids, nucleotides, and sugars taken up by all isolates. However, uptake of specific organic acids, fatty acids and quaternary amines was highly variable across isolates (Supplementary Fig. 11). Profiles of metabolites from the same class, particularly amino acids, nucleotides and aromatic organic acids formed clusters showing similarity in uptake patterns of compounds within the same chemical class across isolates.

We then tested for significant differences in root exudate metabolite uptake across positive and negative responders to root growth. We found that 32 of the 101 metabolites consumed by isolates showed significant (P<0.05) differences across these isolate groups and 13 metabolites had more than 20% difference in metabolite uptake between positive and negative responders (Fig. 6a and Supplementary Fig. 11). The most significant differences (percentage of metabolite depletion from the medium) in substrate preferences among isolates were defined by the cluster of aromatic organic acids (nicotinic, shikimic, salicylic, cinnamic, and indole-3-acetic) (Fig. 6b and Supplementary Fig. 11). Isolates responding positively to the root have up to 48% higher percentage of metabolite depletion from the medium for the organic acids with aromatic rings compared to those isolates that responded negatively to growing roots.

These root exudate components have been shown to influence the composition of rhizosphere microbiomes. Salicylic acid for example, a key regulator of plant metabolism, induces systemic resistance in plants to suppress growth of pathogenic microorganisms[34]. Although some plant pathogens have been reported to degrade it[55], salicylic acid has been shown to be necessary for the assembly of a 'normal' root microbiome of *Arabidopsis thaliana*[16], and together with gamma-aminobutyric acid, salicylic acid concentration has been shown to correlate with specific taxa frequently enriched in the rhizosphere[14]. Taken together with our observations, it appears that the ability to preferentially consume salicylic acid may be a distinguishing feature of rhizosphere bacteria. We also observed that several pentoses showed a higher percentage of uptake by positive responders than by negative responders but conversely we found that nucleosides (cytidine, guanosine, thymidine) were more preferentially consumed by the negative responders (Fig. 6a). Together these data suggest that root exudate chemical composition selectively enriches rhizosphere responders based on their substrate utilization resulting in niche partitioning among soil bacteria in rhizosphere.

Predicting microbial response to root growth based on exudate composition and isolate substrate preferences. We used data on isolate relative abundances in the rhizosphere and their exudate metabolite uptake preferences to build a principal component regression model in order to predict microbial response to plant development (Supplementary Fig. 12). The model identified a strong relationship between exogenous metabolite resources and microbial growth in response to roots. We further extracted the cumulative loading scores from the optimal 11 components, and organic acids (glutaric, nicotinic, indole-3-acetic and threonic acids) were found among the compounds with the highest values (Supplementary Figures 13b and 14a). Despite the limited number of observations used for principal component regression, the predicted top metabolites that drive separation of positive and negative responders using their substrate preferences were in agreement with metabolites identified by other methods in our study (Fig. 6).

This principal component regression model was then used to predict the response of four bacterial isolates to the plant growth based solely on their measured metabolite uptake preferences (Supplementary Fig. 14b). The 16S rRNA gene relative abundances of these four rhizosphere isolates changed in response to plant growth, however these changes were not statistically significant, these isolates were classified as "undefined" (Fig. 2). We used substrate uptake preferences of these isolates as predictors to identify their behavior in rhizosphere. According to the predictive model, *Nocardioides* HA20 revealed a putative negative response to root growth and *Microbacterium* HA36, *Flavobacterium* HB58 and *Cellulomonas* HD24 were identified as positive responders (Supplementary Fig. 14b). Three out of four predicted responses corroborate the putative responses to plant development suggested by their 16S rRNA gene abundance patterns. This suggests that metabolite uptake traits may be particularly valuable predictors of rhizosphere colonization.

Concluding remarks

Interactions between roots, microbes and the soil matrix represents a suite of extremely complex processes. Acknowledging this complexity, the work presented here represents a defined suite of experiments designed to evaluate the potential for metabolic plant-microbial linkages in the rhizosphere of an annual grass in the absence of soil matrix effects. We show that programmed developmental processes in plants result in dynamic patterns of the chemical composition of root exudates. This chemical succession in the rhizosphere interacts with microbial metabolite substrate preferences that can be predicted from genome sequences[32]. We propose that the combination of these plant exudation traits and microbial substrate uptake traits contributes to a metabolic synchronization that underlies microbial community assembly patterns observed in the rhizosphere[2].

Methods

Bacterial isolations and genome sequencing. This study focused on bacterial communities from an annual grassland soil sampled from the University of California Hopland Research and Extension Center (Hopland, CA, USA; 38° 59' 34.5768" N, 123° 4' 3.7704" W). The soil is classified as a coarse-loamy, mesic Ultic Haploxeroll (USDA-NRCS web soil survey; http://websoilsurvey.nrcs.usda.gov) and experiences a Mediterranean-type climate. *Avena barbata* is generally the dominant annual grass present at this field site and is known to have

been a dominant grass in this area for the past century. *Avena fatua* is another wild oat species that widely populates Mediterranean grasslands. Our previous work showed that the rhizosphere microbial communities of these two closely related plants at the same developmental stage are statistically indistinguishable when these species are grown in the same soil[56]. We selected the dominant species, *Avena barbata*, to isolate associated soil bacteria and to measure plant exudation. In order to classify the dynamics of isolated bacteria we used data on previously identified bacterial dynamics in response to *Avena fatua* growing in the same soil. Insignificant differences between the microbiomes of both *Avena* species enabled us to map the dynamics of bacterial isolates over the course of *Avena* rhizosphere development. In this manuscript both of these species will be referred to as *Avena*. Other soil properties are described in Shi et al. (2015)[2]. Growth media for isolation of heterotrophic bacteria were formulated with different concentrations of nutrients, different solidifying agents and anti-oxidant enzymes (R2A 1/10, R2G 1/20, 1/100; OLI; VXylG) according to da Rocha et al., (2015)[57] (Supplementary information). From 289 diverse isolates (Fig. 1), 39 were selected based on phylogeny and relative abundance within the soil to be genome sequenced (Supplementary Figures 1, 2 and 3) using the Illumina HiSeq 2500 platform in accordance with the standard protocols of the DOE Joint Genome Institute (Walnut Creek, CA, USA). Sequences of genomes were deposited, assembled and annotated at the JGI IMG portal (https://img.jgi.doe.gov; Supplementary Data 1, Supplementary Data 2).

16S rRNA phylogeny and successional patterns of isolated bacteria. The successional trajectories of rhizosphere bacterial communities in this soil in response to the growth of *Avena fatua* were previously reported by Shi et al (2015)[2]. Briefly, samples were collected across 12 biological replicates from the rhizosphere of *Avena* and from bulk soil over the course of 0, 3, 6, 9 and 12 weeks of *Avena* growth. These samples were subsequently analyzed by high-throughput sequencing of 16S rRNA genes[2]. To relate the bacterial isolates to the reported bacterial community trajectories, we mapped 16S rRNA sequences of isolates to sequences obtained from bulk and rhizosphere soil during *Avena* growth[2] by comparing the corresponding V4 regions of the isolate 16S rRNA genes (extracted using BLASTN[58] and MUSCLE sequence alignment[59]) with overlapping OTUs defined as having E-values <1e-10 and >= 97% of gene sequence homology. Isolates corresponding to OTUs detected in the soil were assigned the response pattern to plant growth of those OTUs as determined by Shi et al (2015)[2].

Analysis of genomic features of bacterial isolates. We analyzed genome sequences for specific traits related to fitness in the rhizosphere. Traits included those related to growth strategies, substrate uptake and extracellular enzyme production (Supplementary Data 2, Supplementary Figures 3 and 4, Supplementary Table 2). As a proxy for growth strategies, minimum generation times were predicted based on codon usage bias between all genes and a set of highly expressed (ribosomal protein) genes following the linear regression model from Vieira-Silva and Rocha (2010) (Equation 3)[60].

$$(\Delta ENC' = \frac{ENC'_{all} - ENC'_{ribosomal\ protein\ genes}}{ENC'_{all}})$$

Where ENC' is the effective number of codons given G+C composition[61].

For extracellular enzyme traits we calculated gene copy numbers of glycoside hydrolases and auxiliary activity enzymes using a hidden Markov model (HMM) search of protein sequences against the CAZy database[62,63]. For substrate uptake, transporters in the genomes were predicted using an HMM search against TransportDB[64].

The relationships between isolate phylogeny and genome features were determined through analysis of full length 16S rRNA following alignment using MUSCLE 3.8.31[59] and construction of a maximum likelihood tree using FastTree[65] (Supplementary Fig. 3).

***Avena* root exudate collection.** *Avena barbata* (*A. barbata*) wild type seeds were germinated using Milli-Q water and glass wool in the dark at room temperature. Three-day-old seedlings were transferred to six liter hydroponic tubs with half-strength Murashige and Skoog[66] basal salt mixture M524 (Phyto Technology Laboratories, Overland Park, KS). Hydroponic tubs were incubated at 24 °C on 16/8 h day/night cycle, humidity was maintained at 72% and irradiance at 180 μE m^{-2} s^{-1} in growth chambers at the Joint BioEnergy Institute, Emeryville, CA, USA. We refreshed growth medium on a regular basis (every 3 days) to minimize potential microbial growth. Root exudates were collected from 3, 6, 9 and 12 week-old plants corresponding to different developmental stage of the plant: seedling (3 weeks), vegetative (6 and 9 weeks), and senescence (12 weeks) phases respectively. Four biological replicates were collected at each growth stage and exudates were collected from 16 *Avena* plants. To collect root exudates, roots of growing plants were washed in Milli-Q water to remove excess salts from growth media and then transferred to glass cylinders containing 200 ml of sterile Milli-Q water for one hour in a growth chamber[67]. Milli-Q water containing exudates was immediately filter sterilized using 0.22 μm filters (Corning Ink., Corning, NY, USA) and frozen. This exudate collection approach has been used previously for incubation times of up to 2.5 hours without any significant effect of microbial transformation of exudates[38]. The total organic carbon (TOC) concentration in samples was quantified using a Shimadzu TOC-L Analyzer (Shimadzu, Japan) (Supplementary Fig. 5). Exudate samples were lyophilized using a Labconoco FreeZone 2.5 lyophilizer and stored in -80 °C.

Design of bacterial exudate growth medium and cultivation of bacterial isolates for exometabolomics analysis. Sixteen bacterial isolates were selected for metabolite profiling based on their phylogeny, distributions of genomic traits related to rhizosphere carbon utilization and response to plant root growth. The exudate medium was prepared from a base medium (dipotassium phosphate 0.15 g/L, magnesium sulfate 0.012 g/L) to which plant exudates were added. The concentration of exudates added to the growth medium was selected to match the concentration of dissolved organic carbon detected in soil from Hopland, CA (0.025 mg C g^{-1} dry soil) where these isolates originated (Thea Whitman, pers. comm.). When converted to a relevant concentration that bacteria would experience in soil (assuming 20% v/v soil moisture and a soil bulk density of 1.6 g/ml) this corresponds to approximately 125 mg-C/L. Root exudates were first lyophilized, redissolved in water, and then added to growth media to achieve this final C concentration, as determined using a Shimadzu TOC-L Analyzer (Shimadzu, Japan). The resulting medium was subsequently filter sterilized using 0.22 μm filters and inoculated with one of each of the 16 bacterial isolates or incubated as controls. The cultures inoculated at an initial OD_{590}=0.04 and then incubated at 28 °C with shaking at 200 rpm for up to 96 hours. Samples were collected at the early

stationary phase of each isolate, centrifuged, and spent supernatant was frozen at -80 °C. Each isolate culture and uninoculated control samples had four biological replicates (n=68). The concentration of TOC in the uninoculated exudate medium (control samples) and TOC in spent exudate medium from isolate inoculated treatments was quantified (Supplementary Fig. 6). Changes in the TOC concentrations in the spent media from inoculated treatments were compared to that in the uninoculated controls to determine the TOC consumed by each isolate, which was used as an indicator of microbial growth. This revealed that six inoculated samples failed to grow, these samples were excluded from further analysis (Supplementary Fig. 6).

Extraction of metabolites from lyophilized exudates. The TOC of exudates collected at different time points (Supplementary Fig. 5) was measured prior to the sample lyophilization. This information was used to adjust the dilution volumes for the final extracts such that all samples had organic carbon concentrations of approximately 470 mg/L. Here, the necessary volume of cold (-20 °C) methanol containing internal standards 1µg/mL 2-Amino-3-bromo-5-methylbenzoic acid (ABMBA), 5 µg/mL 13C-15N-L-phenylalanine, and 2 µg/mL 9-anthracene carboxylic acid (ACA) was added to the dried exudates and these were sonicated for 30 min using a ultrasonic bath (VWR, Radnor PA). The resulting extracts were filtered using 0.22 µm microcentrifuge PVDF filters (Merck Millipore) and aliquots of 150 µl of methanol extracts were transferred to LC-MS vials for the analysis.

Extraction of metabolites from spent bacterial growth medium. The same starting TOC concentration of growth medium (described above) was used for all treatments (Supplementary Fig. 6). Spent media (1 ml) was collected and then lyophilized and extracted with 150 µl of methanol containing internal standards (as described above). These were then centrifuged at 6000g for one minute and filtered using 0.22 µm microcentrifuge PVDF filters to remove any particles and then analyzed by LC-MS/MS.

Mass spectrometry analysis of exudates and spent bacterial growth media. UHPLC normal phase chromatography was performed using an Agilent 1290 LC stack, with MS and MS/MS data collected using a Q Exactive Orbitrap MS (Thermo Scientific, San Jose, CA). Full MS spectra were collected from *m/z* 70-1050 at 70,000 FWHM resolution, with MS/MS fragmentation data acquired using 10, 20 and 30eV collision energies at 17,500 FWHM resolution. MS instrument parameters included sheath gas flow rate of 50 (arbitrary units, au), auxiliary gas flow rate of 20 (au), sweep gas flow rate of 2 (au), 3 kV spray voltage and 400 °C capillary temperature. Normal phase chromatography was performed using a zic-pHILIC column (Millipore SeQuant ZIC-pHILIC, 150 x 2.1 mm, 5µm, polymeric) at 40 °C at a flow rate of 0.25 mL/min with a 2 µL injection volume. The HILIC column was equilibrated with 100% buffer B (90:10 ACN:H_2O w/ 5mM ammonium acetate) for 1.5 minutes, diluting buffer B down to 50% with buffer A (H_2O w/ 5mM ammonium acetate) 23.5 minutes, down to 40% B over 3.2 minutes, to 0% B over 6.8 minutes, and followed by isocratic elution in 100% buffer A for 3 minutes. Exact mass and retention time coupled with MS/MS fragmentation spectra were used to identify compounds as described below.

Analysis of mass spectrometry data. We first evaluated overall the quality of the dataset by aligning total ion chromatograms from all LC-MS runs using MZmine version 2.26[68] they

demonstrated low variability across biological replicates (Supplementary Figures 7 and 8). Additionally, intensities of the internal standards were assessed from each sample to ensure consistency of signal intensity and retention times from sample-to-sample. Based on the quality control (QC) assessment (mass accuracy within 5 ppm, retention time and peak intensity of the internal standard) three samples that did not pass QC were excluded from further analysis (Supplementary Figure 9). Samples that passed the QC steps were then analyzed using the Metabolite Atlas (https://github.com/biorack/metatlas) for the metabolite feature extraction and annotation[69]. Briefly, metabolites were identified using Metabolite Atlas by matching experimental spectra to our in-house library of authentic standards (accurate mass less than 5 ppm, retention time within 0.5 min and/or match of major MS/MS fragments). In order to maximize the number of compounds identified in our study, in addition to our in-house library of standards we compared the MS/MS data against Metlin[70] and MassBank[71] spectral libraries. Next, MZmine version 2.26[68] was used for manual validation of identified metabolites (accurate mass less than 5 ppm, retention time within 0.5 min and/or match of major MS/MS fragments) to eliminate false identifications. For the metabolites identified in our study (Supplementary Data 3 and 4) we provide a classification of metabolite identification confidence levels recommended by the Metabolomics Standards Initiative Chemical Analysis Working Group of the Metabolomics Society[72]. Briefly, a Level 1 identification, an 'identified' metabolite, requires two independent and orthogonal measures relative to an authentic standard analyzed under the same experimental conditions (for example, *m*/*z* and retention time; *m*/*z* and MS/MS; retention time and MS/MS). 'Putatively' annotated compounds (Level 2) and 'Putatively characterized compound classes' (Level 3) do not have chemical reference standards and annotations are based on spectral similarity to known compounds of a chemical class[72]. Although 87 of the 101 metabolites described in our study were assigned to the Level 1 identifications we also include 12 level three identification (Supplementary Data 3 and 4), which may be of interest. There is ambiguity for many carbohydrates so in these cases we use the following notation: carbohydrate class (standard matched). These data files (Supplementary Data 3 and 4) also provide metabolites peak areas for both positive and negative ionization modes which were used for the statistical analysis of the dynamics of the relative abundances of metabolites across treatments.

In this study we used relative quantification to compare the change in relative abundance of a given metabolite across all samples. A limitation of this commonly used approach is that it does not provide absolute abundances for the various metabolites. Relative comparisons of metabolites were performed by comparison of integrated peak areas in uninoculated control and spent medium from isolate cultivation.

Statistical analyses. To classify the patterns of isolate response to *Avena* growth, we first compared changes in the relative abundance of isolated taxa from 3, 6, 9 and 12 week-old *Avena* plants and bulk soil by using Permutational Analysis of Variance and post-hoc Duncan's multiple range test. The magnitude of change (Δ of isolate abundance) for all isolates was calculated by considering the maximum change in relative abundance from week 0 and any subsequent time point. Responses were classified as Positive, Negative and Undefined according to the sign of the Δ of isolate relative abundance (positive Δ demonstrating an increase in relative abundance over time, negative Δ demonstrating a

decrease in relative abundance over time), undefined indicating no significant change in relative abundance over time (Fig. 2, Supplementary Data 1).

Duncan's multiple range test was also used to test for significant differences in root exudate metabolite composition across the different growth stages of *Avena* (3, 6, 9, 12 week-old plants) (Supplementary Fig. 10). Principal component analysis (PCA) was used to evaluate and visualize relationships between root exudate metabolite profiles at different growth stages of *Avena,* and also to compare the exometabolomes (substrate uptake/release) of bacterial isolates. The Kolmogorov-Smirnov test (KS test) was used to test for statistical differences in the distributions of uptake of specific chemical classes by the classified groups of bacterial isolates (i.e. positive, negative, and undefined responders). Kruskal–Wallis one-way analysis of variance was used to test for statistical differences in the substrate uptake preferences and genomic traits of isolates (Supplementary Fig. 11). All statistical analyses were performed within the R software environment[73] using the "vegan" and "agricolae" packages[74,75].

Principal component regression and prediction of metabolites discriminating bacterial response to root growth. Using exometabolomic data of substrate preferences of bacterial isolates and their calculated abundance dynamics in response to root growth, we tested for discriminatory root exudate metabolites that best predicted the response of these bacterial isolates to growing roots. The overall analysis approach is summarized in Supplementary Fig. 12. Isolates were first classified as positive and negative bacterial rhizosphere responders and enumerated as 1 and 0 respectively; organisms with undefined responses were not included (Supplementary Data 5). Then metabolite uptake data on the biological replicates of all of the isolates were split into training (45 observations) and test (12 observations) data sets. Then PCA was performed and all principal components (PCs) were used to build a principal component regression (PCR) model based on training data in order to predict test set isolate response to root growth based solely on metabolite uptake preferences (Supplementary Data 5). An optimal number of components (n=11), for the PCR model was selected by iterating through extracted PCs to obtain the lowest cross-validation error from 10-fold cross-validation (MSEP < 0.05) (Supplementary Figures 13a and 14a)[76]. Cumulative loading scores across the 11 components were used to identify key discriminatory metabolites whose uptake was predictive of bacterial response to root growth. Using this model we predicted the response of 4 bacterial isolates that experimentally had shown ambiguous (termed 'undefined') responses to root growth. R^2 values of model fit represented the accuracy of model prediction of these isolates' response to root growth. These analyses were performed using packages in the R software environment, including "pls" for principal component regression[76].

Data availability. The genomes of Hopland isolates are publicly available in the Integrated Microbial Genomes (IMG, https://img.jgi.doe.gov) database under IMG study name Mediterranean Grassland Soil Metagenome (MGSM): Enabling a systems view of soil carbon and nitrogen biogeochemistry under a changing climate. All data, including samples that we excluded from the analysis, were deposited to the Global Natural Products Social Molecular Networking (GNPS, https://gnps.ucsd.edu) data repository (*Avena* exudates MSV000081804, Bacterial isolate uptake of exudates MSV000081808, Root exudates components library MSV000081810).

Acknowledgements

This study was supported by the US Department of Energy, Office of Science, Office of Biological Environmental Research including a Genomic Sciences program award no. DE-SC0010570, DE-SC0016247 and DE-SC0014079 to MKF. Work performed at Lawrence Berkeley National Laboratory including DOE Early Career Awards to DL and TRN, and work performed at the DOE Joint Genome Institute (http://www.jgi.gov) and at the DOE Joint BioEnergy Institute (http://www.jbei.org) is supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research through Contract No. DE-AC02-05CH11231. DL was also supported in part by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement number 659910. Isolate genome sequencing was conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, under a Community Science Program award to ELB, supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We thank C. Castanha for background information on soil temperature and are very grateful to our talented former undergraduate students Badamtseteg Jargalsaikhan, Rozina Hossainkhil, Dinh Ly, Severin Ouedraogo and Yen Nguyen for assistance with maintenance of the bacterial isolate collection.

Author contributions

Kateryna Zhalnina[1,2], Katherine B. Louie[1], Zhao Hao[2], Nasim Mansoori[1,3], Ulisses Nunes da Rocha[2,4], Shengjing Shi[5], Heejung Cho[2,6], Ulas Karaoz[2], Dominique Loqué[1,3,6,7], Benjamin P. Bowen[1], Mary K. Firestone[2,8], Trent R. Northen[1], and Eoin L. Brodie[2,8]

KZ, TN, MKF and ELB developed hypotheses. KZ, KL, NM, UNR, SS and DL performed experimental analyses. KZ, HC, UK, ZH, UNR, BB analyzed data. KZ, TN, MKF and ELB wrote the paper. All authors provided comments and edits on the manuscript.

Additional information

Supplementary information is available for this paper. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to TN and ELB.

Competing interests

The authors declare no competing financial interests.

References

1. Hiltner, L. Über neuere Erfahrungen und Probleme auf dem Gebiete der Bodenbakteriologie unter besonderer Berücksichtigung der Gründüngung und Brache. *Arb DLG* **98,** 59–78 (1904).
2. Shi, S. *et al.* Successional Trajectories of Rhizosphere Bacterial Communities over Consecutive Seasons. *mBio* **6,** e00746–15–8 (2015).
3. Chaparro, J. M., Badri, D. V. & Vivanco, J. M. Rhizosphere microbiome assemblage is affected by plant development. *ISME J* **8,** 790–803 (2013).
4. Chaparro, J. M. *et al.* Root Exudation of Phytochemicals in Arabidopsis Follows Specific Patterns That Are Developmentally Programmed and Correlate with Soil Microbial Functions. *PLoS ONE* **8,** e55731–10 (2013).
5. Bulgarelli, D. *et al.* Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* **488,** 91–95 (2012).
6. Grayston, S. J., Wang, S., Campbell, C. D. & Edwards, A. C. Selective influence of plant species on microbial diversity in the rhizosphere. *Soil Biology and Biochemistry* **30,** 369–378 (1998).
7. Peiffer, J. A. *et al.* Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 6548–6553 (2013).
8. DeAngelis, K. M. *et al.* Selective progressive response of soil microbial community to wild oat roots. *ISME J* **3,** 168–178 (2009).
9. Lundberg, D. S. *et al.* Defining the core Arabidopsis thaliana root microbiome. *Nature* **488,** 86–90 (2012).
10. Lu, Y., Rosencrantz, D., Liesack, W. & Conrad, R. Structure and activity of bacterial community inhabiting rice roots and the rhizosphere. *Environmental Microbiology* **8,** 1351–1360 (2006).
11. Bulgarelli, D. *et al.* Structure and Function of the Bacterial Root Microbiota in Wild and Domesticated Barley. *Cell Host & Microbe* **17,** 392–403 (2015).
12. Schreiter, S. *et al.* Effect of the soil type on the microbiome in the rhizosphere of field-grown lettuce. *Front. Microbiol.* **5,** 144 (2014).
13. Pini, F., Galardini, M., Bazzicalupo, M. & Mengoni, A. Plant-bacteria association and symbiosis: are there common genomic traits in alphaproteobacteria? *Genes (Basel)* **2,** 1017–1032 (2011).
14. Badri, D. V., Chaparro, J. M., Zhang, R., Shen, Q. & Vivanco, J. M. Application of Natural Blends of Phytochemicals Derived from the Root Exudates of Arabidopsis to the Soil Reveal That Phenolic-related Compounds Predominantly Modulate the Soil Microbiome. *Journal of Biological Chemistry* **288,** 4502–4512 (2013).
15. Shi, S. *et al.* Effects of selected root exudate components on soil bacterial communities. *FEMS Microbiology Ecology* **77,** 600–610 (2011).
16. Lebeis, S. L. *et al.* PLANT MICROBIOME. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* **349,** 860–864 (2015).

17. Lynch, J. M. & Whipps, J. M. Substrate flow in the rhizosphere. *Plant Soil* **129,** 1–10 (1990).
18. Badri, D. V. & Vivanco, J. M. Regulation and function of root exudates. *Plant, Cell & Environment* **32,** 666–681 (2009).
19. Baetz, U. & Martinoia, E. Root exudates: the hidden part of plant defense. *Trends in Plant Science* **19,** 90–98 (2014).
20. Bais, H. P., Weir, T. L., Perry, L. G., Gilroy, S. & Vivanco, J. M. The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu. Rev. Plant Biol.* **57,** 233–266 (2006).
21. Aulakh, M. S., Wassmann, R., Bueno, C., Kreuzwieser, J. & Rennenberg, H. Characterization of Root Exudates at Different Growth Stages of Ten Rice (Oryza sativa L.) Cultivars. *Plant biology* **3,** 139–148 (2001).
22. Jones, D. L. Organic acids in rhizosphere - a critical review. *Plant Soil* **205,** 25–44 (1998).
23. Iannucci, A., Fragasso, M., Platani, C. & Papa, R. Plant growth and phenolic compounds in the rhizosphere soil of wild oat (Avena fatua L.). *Front. Plant Sci.* **4,** (2013).
24. Dakora, F. D. & Phillips, D. A. in *Food Security in Nutrient-Stressed Environments: Exploiting Plants' Genetic Capabilities* 201–213 (Springer Netherlands, 2002). doi:10.1007/978-94-017-1570-6_23
25. Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E. V. L. & Schulze-Lefert, P. Structure and Functions of the Bacterial Microbiota of Plants. *Annu. Rev. Plant Biol.* **64,** 807–838 (2013).
26. Hassan, S. & Mathesius, U. The role of flavonoids in root-rhizosphere signalling: opportunities and challenges for improving plant-microbe interactions. *Journal of Experimental Botany* **63,** 3429–3444 (2012).
27. Corral-Lugo, A., Daddaoua, A., Ortega, A., Espinosa-Urgel, M. & Krell, T. Rosmarinic acid is a homoserine lactone mimic produced by plants that activates a bacterial quorum-sensing regulator. *Sci Signal* **9,** ra1–ra1 (2016).
28. Huang, X.-F. *et al.* Rhizosphere interactions: root exudates, microbes, and microbial communities 1. *Botany* **92,** 267–275 (2014).
29. Beauregard, P. B., Chai, Y., Vlamakis, H., Losick, R. & Kolter, R. Bacillus subtilis biofilm induction by plant polysaccharides. *Proc. Natl. Acad. Sci. U.S.A.* **110,** E1621–30 (2013).
30. Cai, T. *et al.* Host legume-exuded antimetabolites optimize the symbiotic rhizosphere. *Mol. Microbiol.* **73,** 507–517 (2009).
31. Silva, L. P. & Northen, T. R. Exometabolomics and MSI: deconstructing how cells interact to transform their small molecule environment. *Current Opinion in Biotechnology* **34,** 209–216 (2015).
32. Baran, R. *et al.* Exometabolite niche partitioning among sympatric soil bacteria. *Nat Commun* **6,** 8289 (2015).
33. Mao, Y., Li, X., Smyth, E. M., Yannarell, A. C. & Mackie, R. I. Enrichment of specific bacterial and eukaryotic microbes in the rhizosphere of switchgrass (Panicum virgatum L.) through root exudates. *Environmental Microbiology Reports* **6,** 293–306 (2014).

34. Chisholm, S. T., Coaker, G., Day, B. & Staskawicz, B. J. Host-Microbe Interactions: Shaping the Evolution of the Plant Immune Response. *Cell* **124,** 803–814 (2006).
35. Souza, R. de, Ambrosini, A. & Passaglia, L. M. P. Plant growth-promoting bacteria as inoculants in agricultural soils. *Genet. Mol. Biol.* **38,** 401–419 (2015).
36. Lambers, H., Mougel, C., Jaillard, B. & Hinsinger, P. Plant-microbe-soil interactions in the rhizosphere: an evolutionary perspective. *Plant Soil* **321,** 83–115 (2009).
37. Roller, B. R. K., Stoddard, S. F. & Schmidt, T. M. Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nat. Microbiol* **1,** 1–7 (2016).
38. Neumann, G. & Römheld, V. in *The Rhizosphere* **20072634,** 23–72 (CRC Press, 2009).
39. Walker, T. S., Bais, H. P., Grotewold, E. & Vivanco, J. M. Root exudation and rhizosphere biology. *PLANT PHYSIOLOGY* **132,** 44–51 (2003).
40. Salerno, G. Origin of sucrose metabolism in higher plants: when, how and why? *Trends in Plant Science* **8,** 63–69 (2003).
41. Jaeger, C., Lindow, S., Miller, W., Clark, E. & Firestone, M. Mapping of sugar and amino acid availability in soil around roots with bacterial sensors of sucrose and tryptophan. *Applied and Environmental Microbiology* **65,** 2685–2690 (1999).
42. Vargas, W. A., Mandawe, J. C. & Kenerley, C. M. Plant-Derived Sucrose Is a Key Element in the Symbiotic Association between Trichoderma virens and Maize Plants. *PLANT PHYSIOLOGY* **151,** 792–808 (2009).
43. Lemoine, R. *et al.* Source-to-sink transport of sugar and regulation by environmental factors. *Front. Plant Sci.* **4,** 1–21 (2013).
44. Lattanzio, V. & Lattanzio, V. Role of phenolics in the resistance mechanisms of plants against fungal pathogens and insects. *... : Advances in research* (2006).
45. Iqbal, N., Nazar, R. & Khan, N. A. *Osmolytes and Plants Acclimation to Changing Environment: Emerging Omics Technologies*. (2015).
46. Noodén, L. D. in *Senescence and Aging in Plants* 329–368 (Elsevier, 1988). doi:10.1016/B978-0-12-520920-5.50016-X
47. Brown, J. H., Paliyath, G. & Thompson, J. E. *Brown: Physiological mechanisms of plant senescence - Google Scholar*. (Plant physiology, 1991). doi:10.1111/j.1399-3054.1981.tb08516.x/full
48. Mueller-Roeber, B. & Balazadeh, S. Auxin and Its Role in Plant Senescence. *J Plant Growth Regul* **33,** 21–33 (2013).
49. Pilet, P.-E. & Saugy, M. Effect on Root Growth of Endogenous and Applied IAA and ABA A Critical Reexamination. *PLANT PHYSIOLOGY* **83,** 33–38 (1987).
50. Ramachandran, V. K., East, A. K., Karunakaran, R., Downie, J. A. & Poole, P. S. Adaptation of Rhizobium leguminosarum to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics. *Genome Biol.* **12,** R106 (2011).

51. Lugtenberg, B. J. J., Kravchenko, L. V. & Simons, M. Tomato seed and root exudate sugars: composition, utilization by Pseudomonas biocontrol strains and role in rhizosphere colonization. *Environmental Microbiology* **1,** 439–446 (1999).
52. Lugtenberg, B. J., Dekkers, L. & Bloemberg, G. V. Molecular determinants of rhizosphere colonization by Pseudomonas. *Annu Rev Phytopathol* **39,** 461–490 (2001).
53. Blakley, E. R. & Simpson, F. J. The microbial metabolism of cinnamic acid. *Canadian Journal of Microbiology* **10,** 175–185 (1964).
54. Yuan, J. *et al.* Organic acids from root exudates of banana help root colonization of PGPR strain Bacillus amyloliquefaciens NJN-6. *Nature Publishing Group* 1–8 (2015). doi:10.1038/srep13438
55. Lowe-Power, T. M. *et al.* Degradation of the Plant Defense Signal Salicylic Acid Protects Ralstonia solanacearum from Toxicity and Enhances Virulence on Tobacco. *mBio* **7,** e00656–16 (2016).
56. Nuccio, E. E. *et al.* Climate and edaphic controllers influence rhizosphere community assembly for a wild annual grass. *Ecology* **97,** 1307–1318 (2016).
57. Nunes da Rocha, U. *et al.* Isolation of a significant fraction of non-phototroph diversity from a desert Biological Soil Crust. *Front. Microbiol.* **6,** 277 (2015).
58. McGinnis, S. & Madden, T. L. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Research* **32,** W20–5 (2004).
59. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32,** 1792–1797 (2004).
60. Vieira-Silva, S. & Rocha, E. P. C. The systemic imprint of growth and its uses in ecological (meta)genomics. *PLoS Genet.* **6,** e1000808 (2010).
61. Subramanian, S. Nearly Neutrality and the Evolution of Codon Usage Bias in Eukaryotic Genomes. *Genetics* **178,** 2429–2432 (2008).
62. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research* **42,** D490–5 (2014).
63. Yin, C., Li, S. & Li, Q. Network traffic classification via HMM under the guidance of syntactic structure. *Computer Networks* **56,** 1814–1825 (2012).
64. Ren, Q., Chen, K. & Paulsen, I. T. TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. *Nucleic Acids Research* **35,** D274–9 (2007).
65. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5,** e9490 (2010).
66. Murashige, T. & Skoog, F. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum* **15,** 473–497 (1962).
67. Khorassani, R. *et al.* Citramalic acid and salicylic acid in sugar beet root exudates solubilize soil phosphorus. *BMC Plant Biology* **11,** 121 (2011).
68. Pluskal, T., Castillo, S., Villar-Briones, A. & Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11,** 395 (2010).

69. Bowen, B. P. & Northen, T. R. Dealing with the unknown: metabolomics and metabolite atlases. *J. Am. Soc. Mass Spectrom.* **21,** 1471–1476 (2010).
70. Smith, C. A. *et al.* METLIN. *Therapeutic Drug Monitoring* **27,** 747–751 (2005).
71. Horai, H. *et al.* MassBank: a public repository for sharing mass spectral data for life sciences. *Journal of Mass Spectrometry* **45,** 703–714 (2010).
72. Sumner, L. W. *et al.* Proposed minimum reporting standards for chemical analysis. *Metabolomics* **3,** 211–221 (2007).
73. Team, R. C. R: A Language and Environment for Statistical Computing. *www.R-project.org* (2017). Available at: www.R-project.org.
74. Blanchet, F. G. *et al.* vegan: Community Ecology Package. R package version 2.4-2. *https://CRAN.R-project.org/package=vegan* (2017).
75. de Mendiburu, F. agricolae: Statistical Procedures for Agricultural Research. R package version 1.2-4. *https://CRAN.R-project.org/package=agricolae* (2016).
76. Mevik, B.-H. & Wehrens, R. The plsPackage: Principal Component and Partial Least Squares Regression in R. *Journal of Statistical Software* **18,** 1–23 (2007).

Figures and legends

Figure 1 | Cladogram showing phylogenetic relationships between 289 soil heterotrophic bacterial isolates, and their origin (media type). Leaf labels indicate representative sequence IDs. Rings, from the inner to the outside circles, represent: (1) genome-sequenced isolates (black blocks); (2) class level taxonomy of isolate; (3) medium on which isolate was originally obtained.

Figure 2 | Growth response in soil of bacterial isolates to *Avena* growth based on changes in 16S rRNA gene abundance. Each bar represents the change (Δ) of isolate abundance between bulk soil at week 0 and the point of maximum change over the developmental stages of *Avena*. Δ of isolate abundance was normalized by the total number of 16S rRNA reads identified for each isolate over all developmental stages of *Avena*. Positive responders (n=19), negative responders (n=8), and * indicates undefined responders (n=11) with a non-significant response relative to week 0 bulk soil relative abundance (* P >0.05, Permutational Analysis of Variance and post-hoc Duncan's multiple range test). Bullet points indicate isolates selected for exometabolite profiling.

Figure 3 | Distributions of select traits in the genomes of soil bacterial isolates classified into two groups based on the response to plant root growth. Positive responders (n=19), negative responders (n=8). **a**, Minimum generation times predicted from genome sequences. **b**, Genome size of isolates. **c**, Extracellular enzymes for plant polymer degradation. **d**, Monomer transporters. All gene frequencies were adjusted for differences in genome size. Total number of transporters shown as percent of transporters per genome. In each boxplot, a point denotes a single metabolic trait or a single gene. The top and bottom of each box represent the 25th and 75th percentiles, the horizontal line inside each box represents the 50th percentile/median and the whiskers represent the range of points excluding outliers. Outliers denoted as large points outside whiskers. Differences in the distributions of traits between the two groups of isolates were evaluated using the Kruskal–Wallis one-way analysis of variance and traits with significant differences (P<0.05) are denoted by *.

Figure 4 | Changes in *A. barbata* exudation through plant development (weeks 3, 6, 9, 12). a, Principal component analysis of the exudate profiles of *A. barbata* at each time point (n=16 exudate profiles). **b**, Changes in abundance of each exudate compound across plant developmental stages (n=16 exudate profiles). All abundances were normalized to a range between 0 and 1, with 1 representing sum of each metabolite released. Metabolites with the highest abundance at each time point are highlighted with the same color. Duncan's multiple range test was used to determine compounds with significant differences across different developmental stages (P<0.05).

Figure 5 | Distributions of root exudate metabolite uptake by isolates presented as percent of uptake from the exudate medium. Bar colors represent positive (green) and negative (purple) responders to root growth (n=12 isolates). In each boxplot, a point denotes a single metabolite and its percent uptake in a single incubation. Diamond symbols in each

boxplot represent the mean. The box boundaries represent the first and third quartiles of the distribution and the median is represented as the horizontal line inside each box. Box plot whiskers span 1.5 times the interquartile range of the distribution. A significant difference between distributions were determined using the Kolmogorov-Smirnov test and are denoted with **(P<1e-08) and *(P<0.01).

Figure 6 | Substrate preferences of positive and negative responders. a, Metabolites with significant (P<0.05) differences in their uptake from *Avena* exudates by isolates (n=12 isolates) (Kruskal-Wallis test). **b**, Uptake of aromatic organic acids from exudate media by isolates with positive (n=8 isolates) and negative (n=4 isolates) responses to plant growth. In each barplot, a point denotes a biological replicate measurement of percent uptake of aromatic acid by an isolate. Metabolite uptake is given as a percentage of metabolite depletion by isolate from the medium, and error bars show standard error of the mean.

Legend:

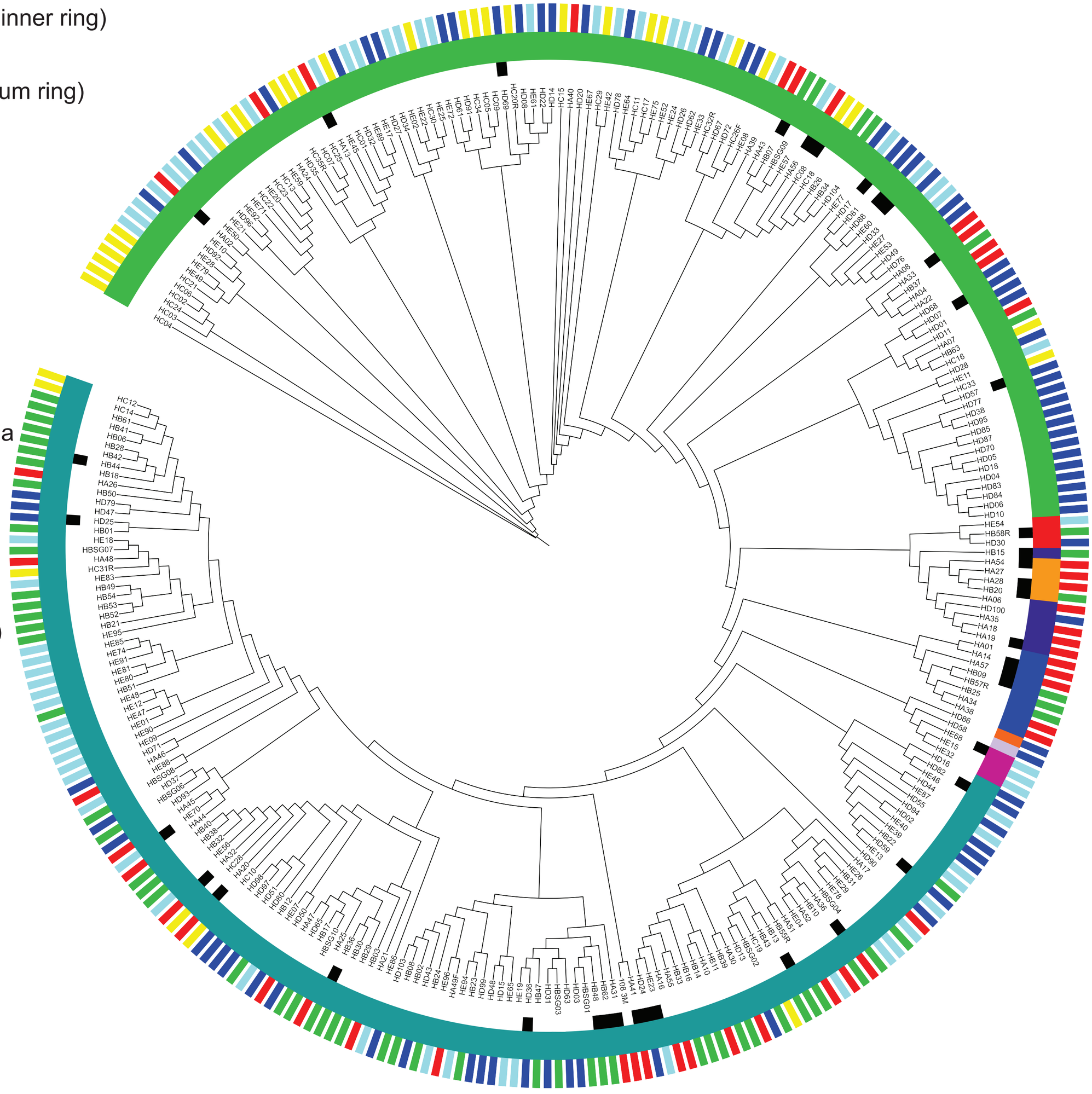
Sequenced isolates (inner ring)

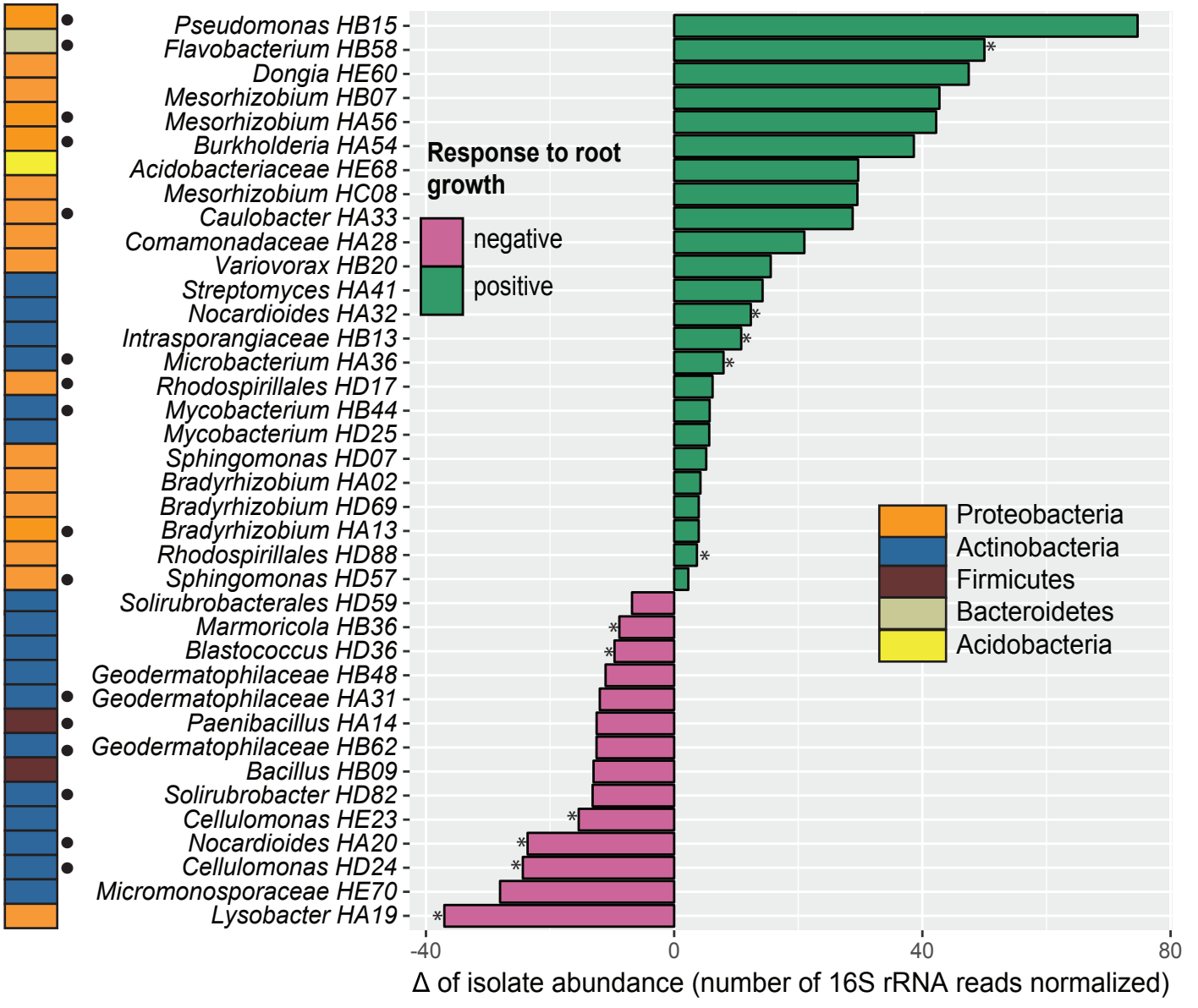
Dataset Phylogeny (medium ring)

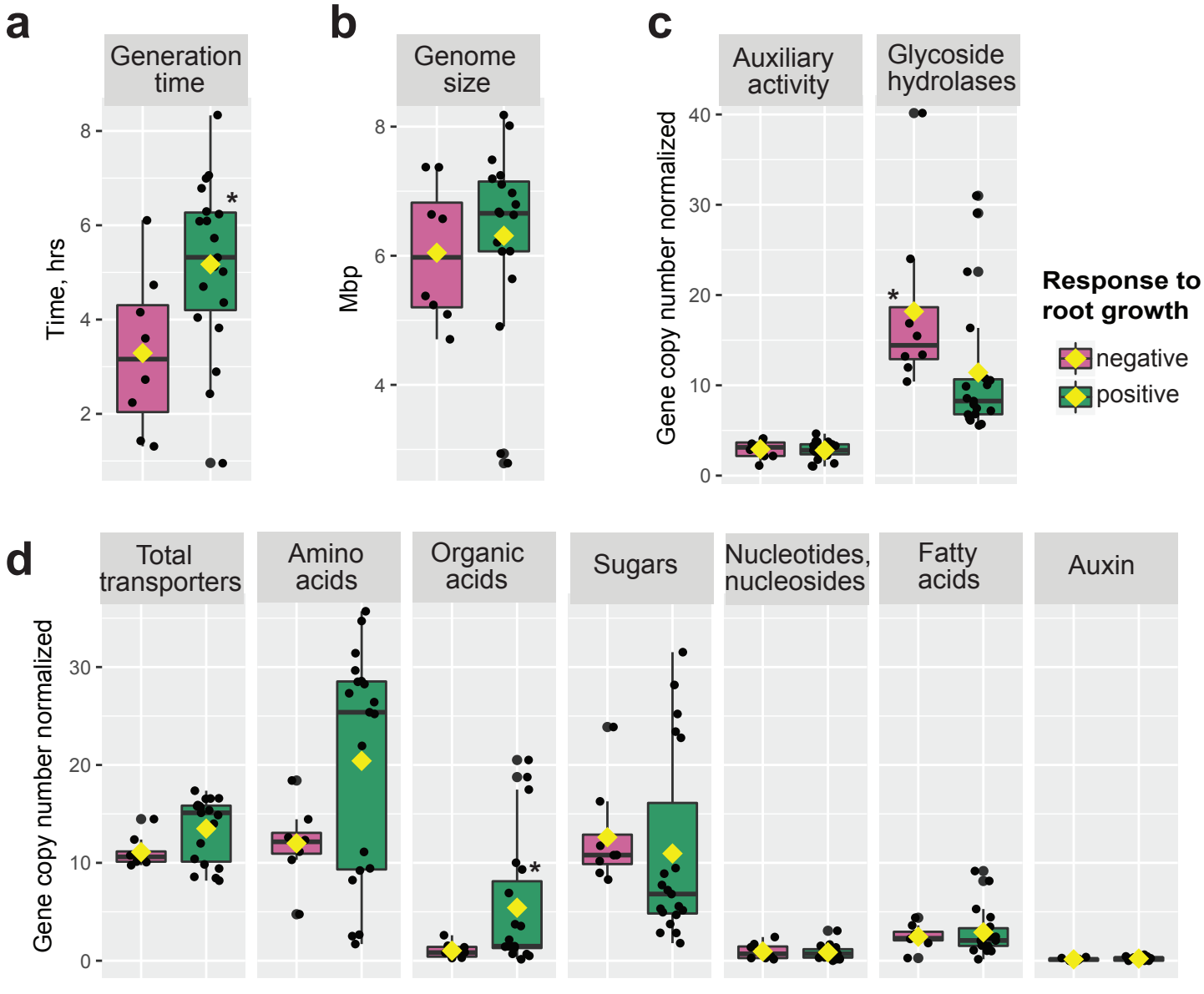
- Acidobacteria
- Actinobacteria
- Alphaproteobacteria
- Betaproteobacteria
- Bacteroidetes
- Firmicutes
- Gammaproteobacteria
- Gemmatimonadetes
- Verrucomicrobia

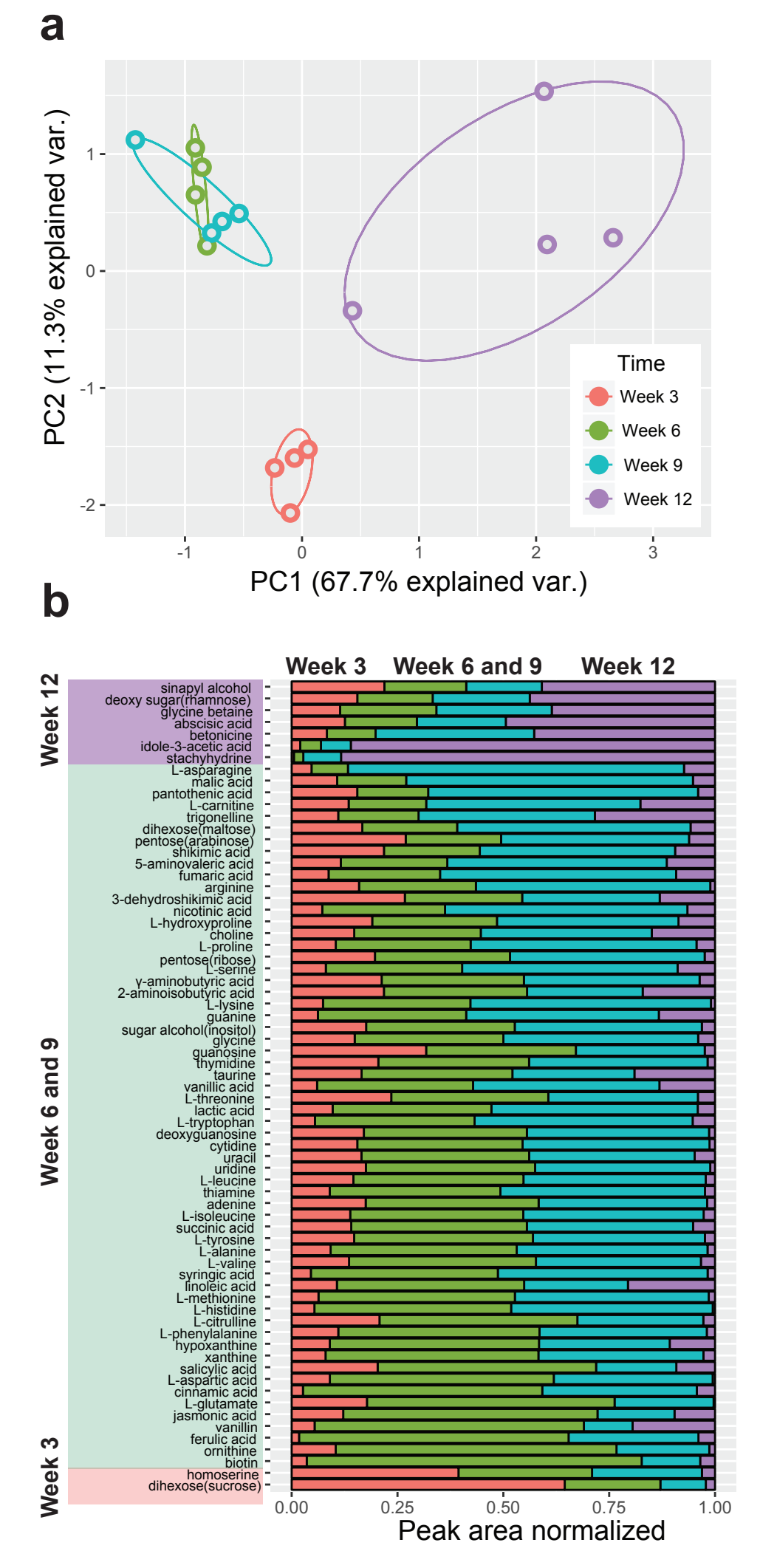
Dataset Media (outer ring)

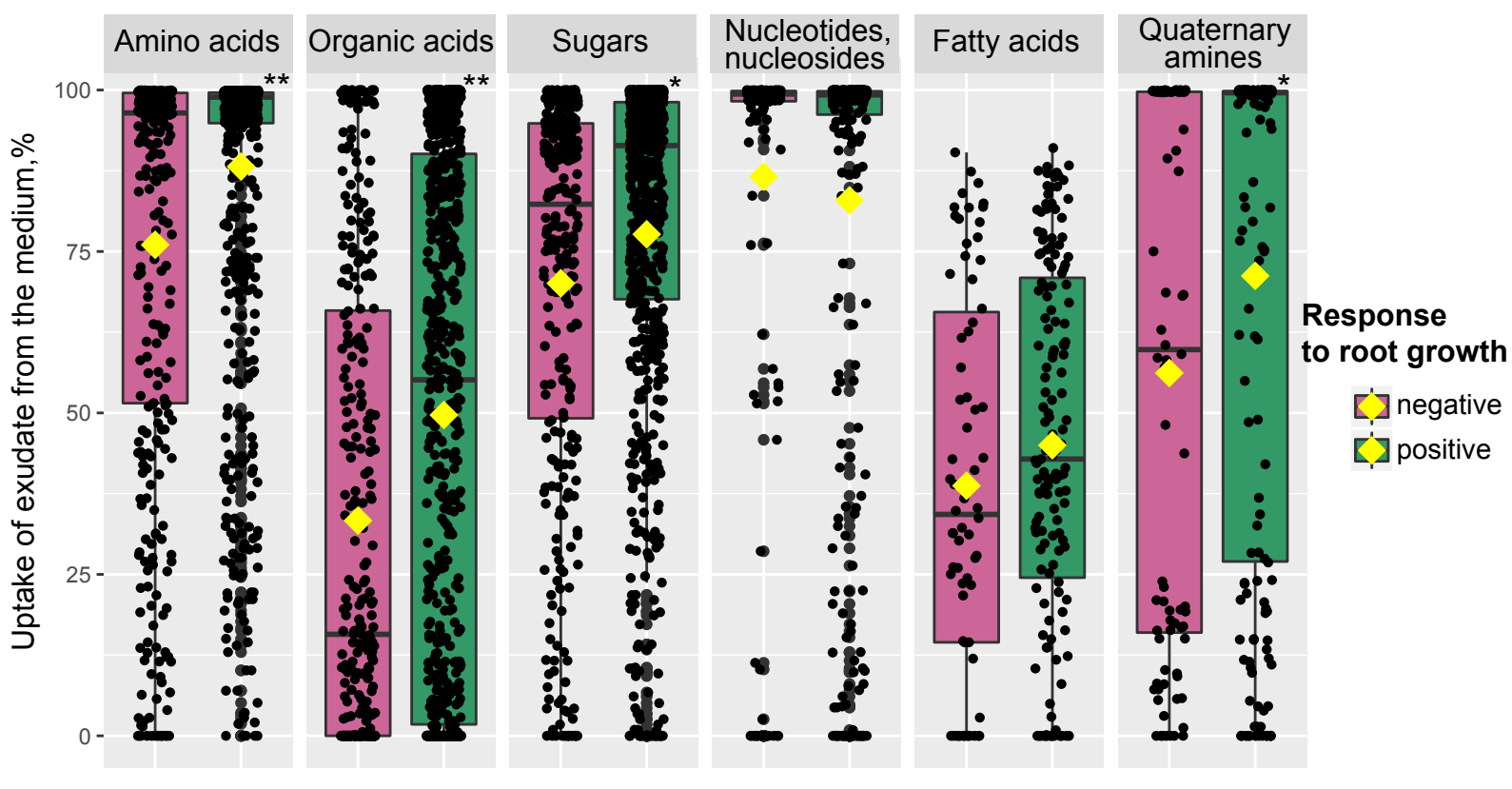
- R2A 1/10
- R2A 1/20
- R2A 1/100
- OLI
- VXylG



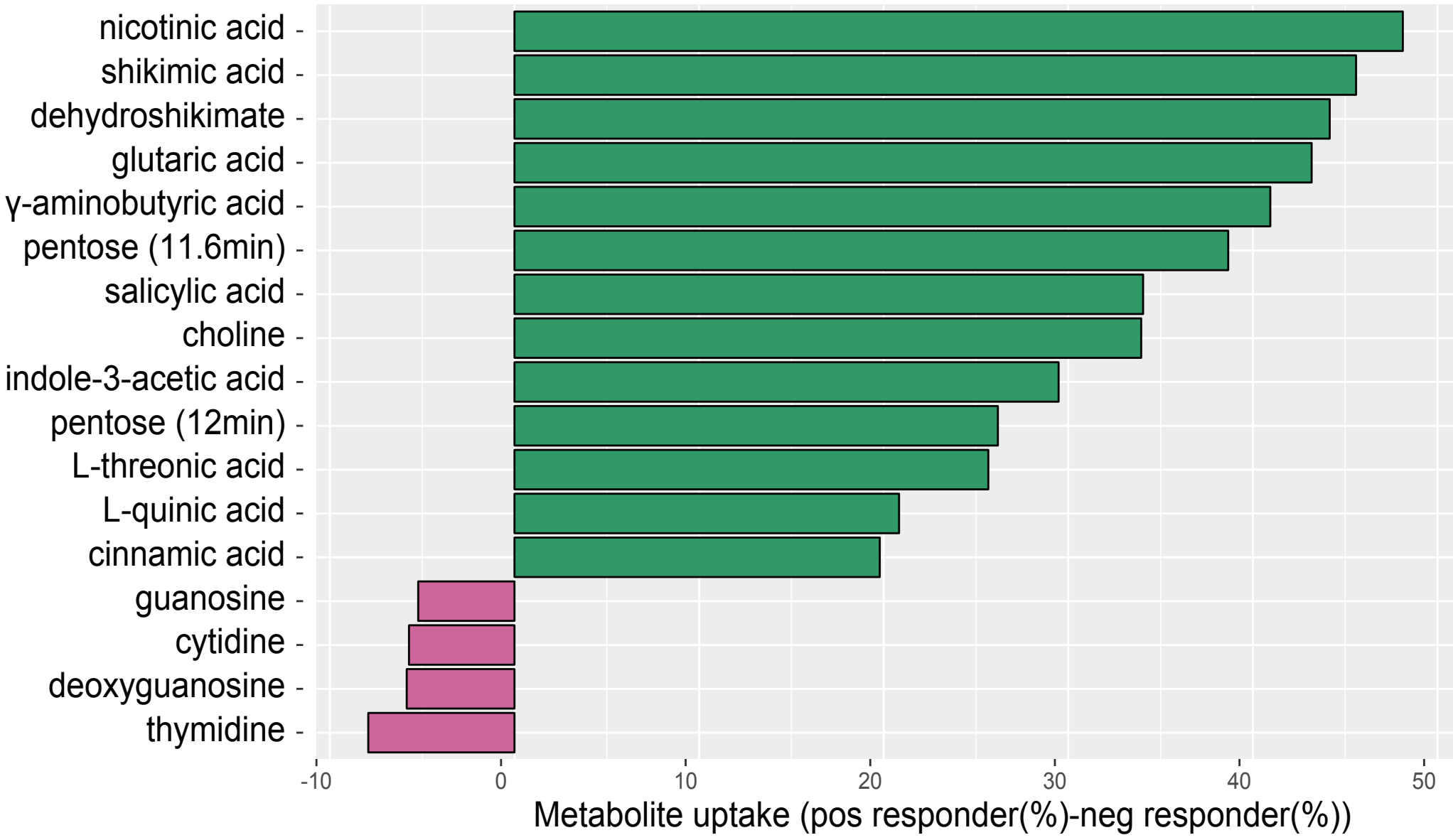

Pseudomonas HB15
Flavobacterium HB58
Dongia HE60
Mesorhizobium HB07
Mesorhizobium HA56
Burkholderia HA54
Acidobacteriaceae HE68
Mesorhizobium HC08
Caulobacter HA33
Comamonadaceae HA28
Variovorax HB20
Streptomyces HA41
Nocardioides HA32
Intrasporangiaceae HB13
Microbacterium HA36
Rhodospirillales HD17
Mycobacterium HB44
Mycobacterium HD25
Sphingomonas HD07
Bradyrhizobium HA02
Bradyrhizobium HD69
Bradyrhizobium HA13
Rhodospirillales HD88
Sphingomonas HD57
Solirubrobacterales HD59
Marmoricola HB36
Blastococcus HD36
Geodermatophilaceae HB48
Geodermatophilaceae HA31
Paenibacillus HA14
Geodermatophilaceae HB62
Bacillus HB09
Solirubrobacter HD82
Cellulomonas HE23
Nocardioides HA20
Cellulomonas HD24
Micromonosporaceae HE70
Lysobacter HA19
Response to root growth
negative
positive
Proteobacteria
Actinobacteria
Firmicutes
Bacteroidetes
Acidobacteria
-40
0
40
80
Δ of isolate abundance (number of 16S rRNA reads normalized)


a
Generation time
Time, hrs
8
6
4
2
*
b
Genome size
Mbp
8
6
4
c
Auxiliary activity
Glycoside hydrolases
Gene copy number normalized
40
30
20
10
0
*
Response to root growth
negative
positive
d
Total transporters
Amino acids
Organic acids
Sugars
Nucleotides, nucleosides
Fatty acids
Auxin
Gene copy number normalized
30
20
10
0
*

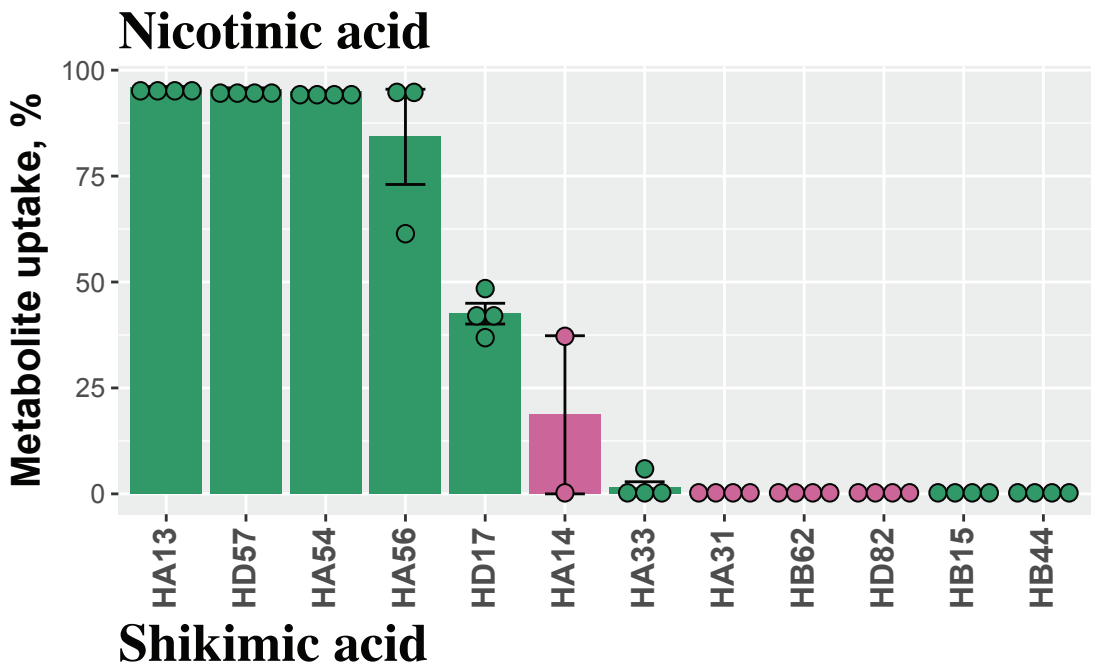

a
PC2 (11.3% explained var.)
PC1 (67.7% explained var.)
Time
Week 3
Week 6
Week 9
Week 12
b
Week 3
Week 6 and 9
Week 12
Peak area normalized
sinapyl alcohol
deoxy sugar(rhamnose)
glycine betaine
abscisic acid
betonicine
idole-3-acetic acid
stachyhydrine
L-asparagine
malic acid
pantothenic acid
L-carnitine
trigonelline
dihexose(maltose)
pentose(arabinose)
shikimic acid
5-aminovaleric acid
fumaric acid
arginine
3-dehydroshikimic acid
nicotinic acid
L-hydroxyproline
choline
L-proline
pentose(ribose)
L-serine
γ-aminobutyric acid
2-aminoisobutyric acid
L-lysine
guanine
sugar alcohol(inositol)
glycine
guanosine
thymidine
taurine
vanillic acid
L-threonine
lactic acid
L-tryptophan
deoxyguanosine
cytidine
uracil
uridine
L-leucine
thiamine
adenine
L-isoleucine
succinic acid
L-tyrosine
L-alanine
L-valine
syringic acid
linoleic acid
L-methionine
L-histidine
L-citrulline
L-phenylalanine
hypoxanthine
xanthine
salicylic acid
L-aspartic acid
cinnamic acid
L-glutamate
jasmonic acid
vanillin
ferulic acid
ornithine
biotin
homoserine
dihexose(sucrose)

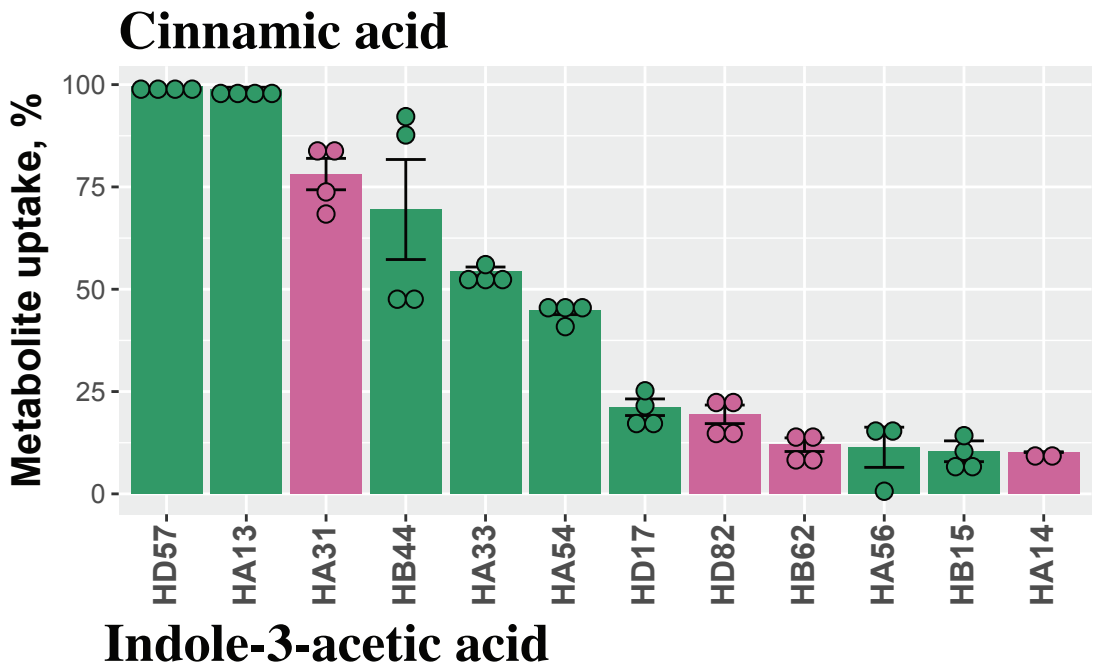

Amino acids
Organic acids
Sugars
Nucleotides, nucleosides
Fatty acids
Quaternary amines
**
**
*
*
Uptake of exudate from the medium,%
100
75
50
25
0
Response to root growth
negative
positive

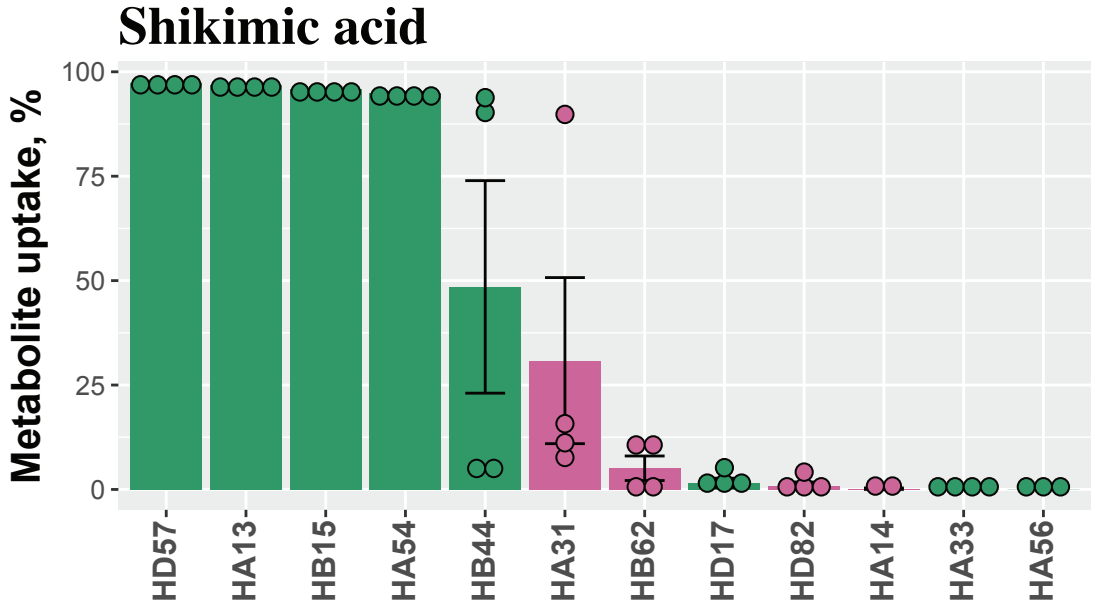
a

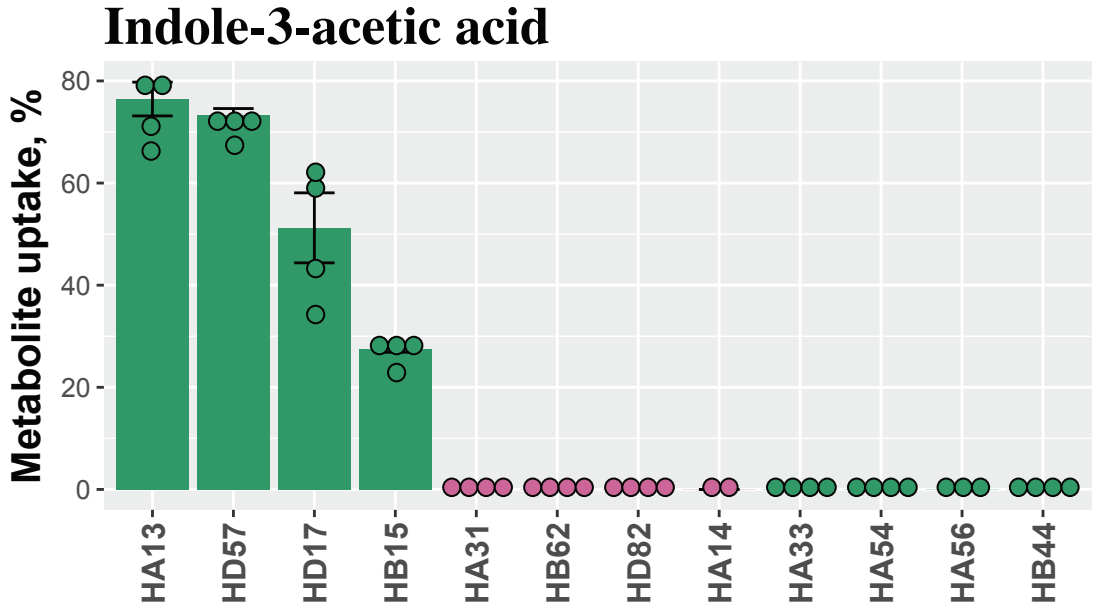

nicotinic acid
shikimic acid
dehydroshikimate
glutaric acid
γ-aminobutyric acid
pentose (11.6min)
salicylic acid
choline
indole-3-acetic acid
pentose (12min)
L-threonic acid
L-quinic acid
cinnamic acid
guanosine
cytidine
deoxyguanosine
thymidine
-10
0
10
20
30
40
50
Metabolite uptake (pos responder(%)-neg responder(%))

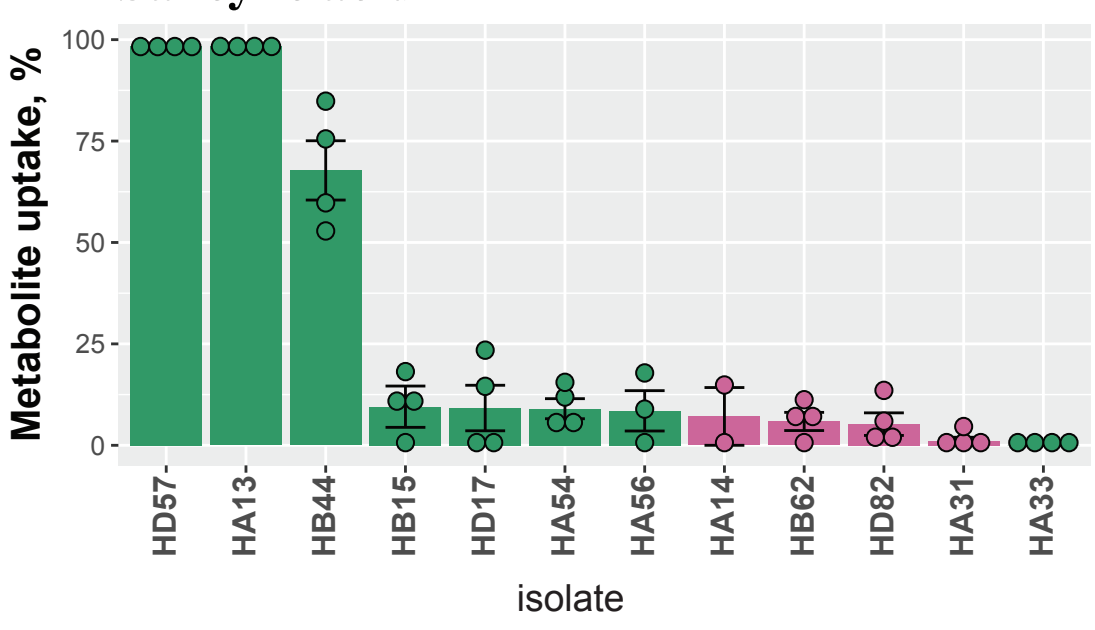
b

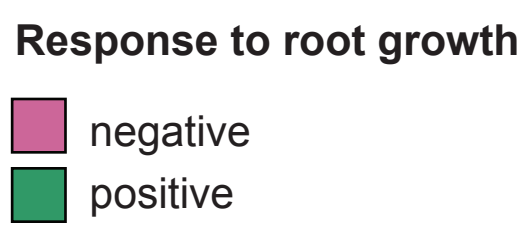

Nicotinic acid
Metabolite uptake, %
100
75
50
25
0
HA13
HD57
HA54
HA56
HD17
HA14
HA33
HA31
HB62
HD82
HB15
HB44


Cinnamic acid
Metabolite uptake, %
100
75
50
25
0
HD57
HA13
HA31
HB44
HA33
HA54
HD17
HD82
HB62
HA56
HB15
HA14


Shikimic acid
Metabolite uptake, %
100
75
50
25
0
HD57
HA13
HB15
HA54
HB44
HA31
HB62
HD17
HD82
HA14
HA33
HA56


Indole-3-acetic acid
Metabolite uptake, %
80
60
40
20
0
HA13
HD57
HD17
HB15
HA31
HB62
HD82
HA14
HA33
HA54
HA56
HB44
isolate


Salicylic acid
Metabolite uptake, %
100
75
50
25
0
HD57
HA13
HB44
HB15
HD17
HA54
HA56
HA14
HB62
HD82
HA31
HA33
isolate


Response to root growth
negative
positive