nature microbiology

ARTICLES

https://doi.org/10.1038/s41564-018-0129-3

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*]

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.

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, geographical 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 the number of bacteria of the Alphaproteobacteria subphylum was reported in rhizosphere soil of various 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 defence, 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 physical and chemical environment of the soil. However, despite the growing number of studies demonstrating that plant development influences the composition of the 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, 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 or static property and varies depending on plant species, developmental stage, root traits, environmental conditions, nutrition and soil type, among other factors[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 effect of small signalling molecules (including acyl-homoserine lactones, flavonoids and non-proteinogenic amino acids, among others)[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 combines

[1]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [2]Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [3]Joint BioEnergy Institute, Biosystems Engineering Division, Lawrence Berkeley National Laboratory, Emeryville, CA, USA. [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, CA, USA. [7]INSA de Lyon, CNRS, UMR5240, Microbiologie, Adaptation et Pathogénie, Université Claude Bernard Lyon 1, Villeurbanne, France. [8]Department of Environmental Science, Policy and Management, University of California, Berkeley, CA, USA. *e-mail: trnorthen@lbl.gov; elbrodie@lbl.gov

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, antioxidants, 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 Alphaproteobacteria that are 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 Figs. 1–3 and Supplementary Data 1). Non-redundant operational taxonomic units (OTUs) matched to these 39 isolates together represent approximately 10–12% of the total bacterial community in

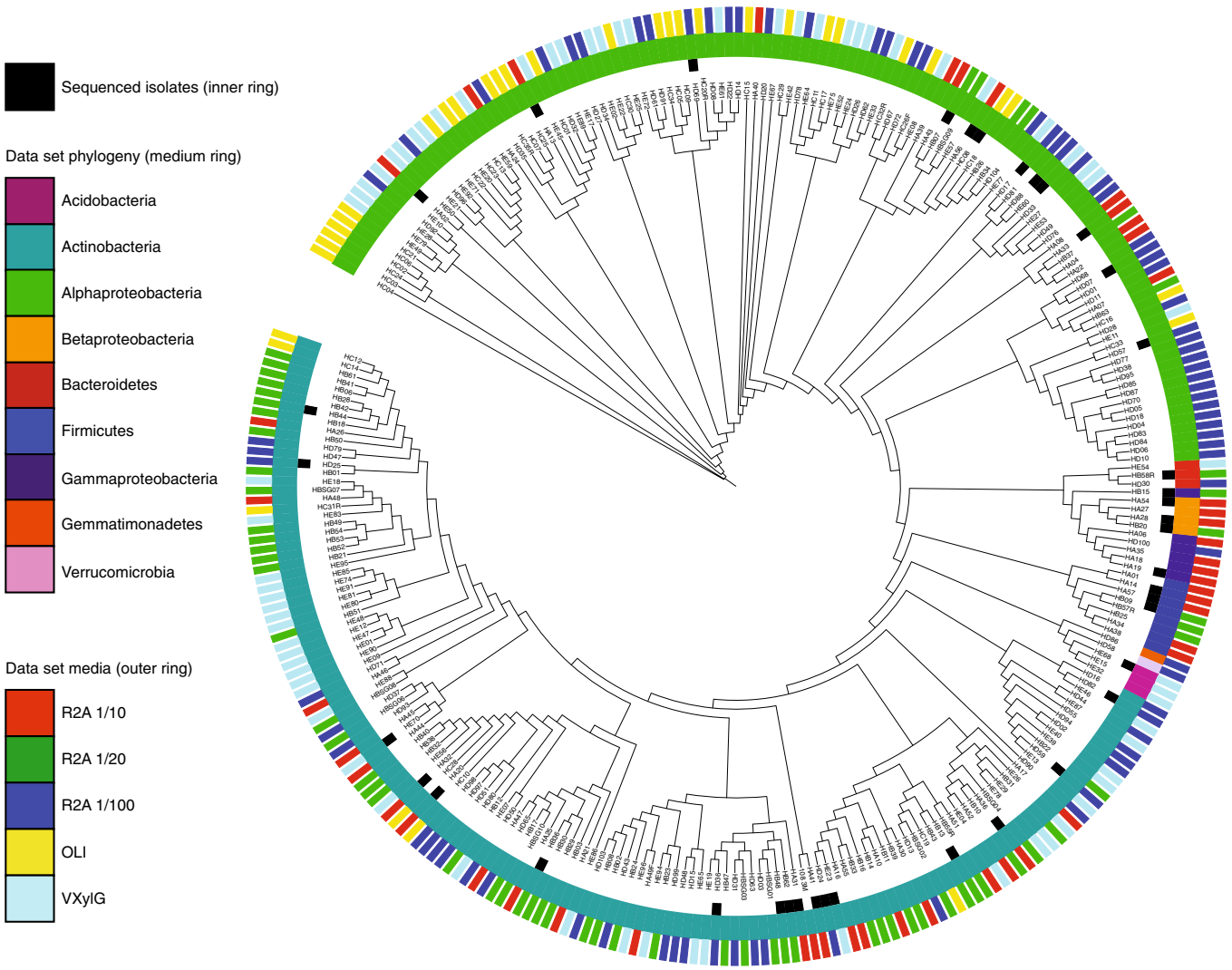

Fig. 1 | Cladogram showing phylogenetic relationships between 289 soil heterotrophic bacterial isolates and their origin (media type). Leaf labels indicate representative sequence IDs. Rings, from the inner to the outside circles, represent: (1) genome-sequenced isolates (black blocks); (2) class-level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

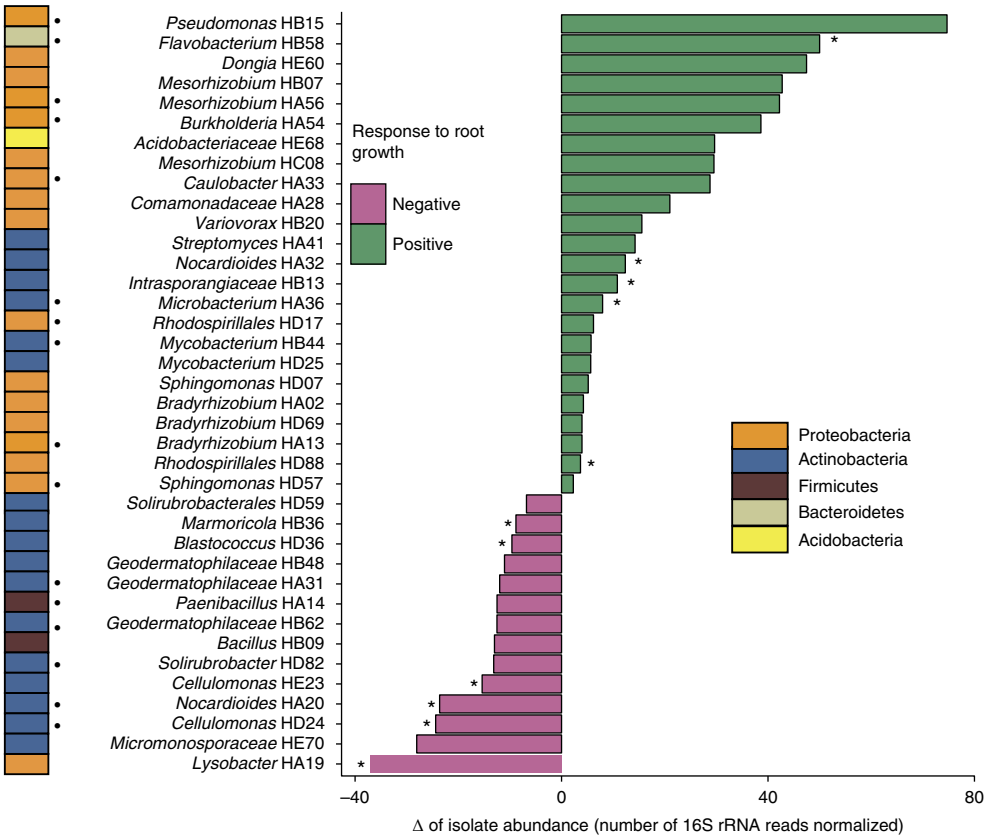

Fig. 2 | Growth response in the soil of bacterial isolates to *Avena* growth based on changes in 16S rRNA gene abundance. Each bar represents the change (Δ) of isolate abundance between bulk soil at week 0 and the point of maximum change over the developmental stages of *Avena*. Δ of isolate abundance 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 undefined responders (denoted by *) 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; see Supplementary Data 1 for details): $n=11$. Black dots indicate isolates that were selected for exometabolite profiling.

this rhizosphere soil, and 17 of these isolates had a relative abundance of 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 ribosomal RNA 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 comprised isolates related to the Proteobacteria (*Dongia*, *Rhodospirillales*, *Sphingomonas*, *Mesorhizobium*, *Bradyrhizobium*, *Caulobacter*, *Burkholderia*, *Variovorax* and *Pseudomonas*) and a limited number of the Actinobacteria (*Mycobacterium* and *Streptomyces*). Isolates from the negative-response group were mostly from the Actinobacteria and Firmicutes (*Bacillus* and *Paenibacillus*). In general, the relative abundance of Actinobacteria declined, whereas the relative abundance of Proteobacteria, particularly alphaproteobacteria, increased during plant growth. This observation is consistent with many other successional studies: including *Arabidopsis thaliana*[4], wheat[11], rice[10], switchgrass[33] and 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 that were 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 that are favourable to life in the rhizosphere[36]. To identify these traits, we analysed their genomes for the presence of features that are hypothesized to be important for both rhizosphere growth and soil organic matter transformation (Fig. 3a–d, Supplementary Fig. 3 and Supplementary Data 2). Specifically, we focused on traits associated with the acquisition of carbon substrates, such as macromolecule depolymerization enzymes and monomer transport, in addition to predicted generation times. We analysed 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

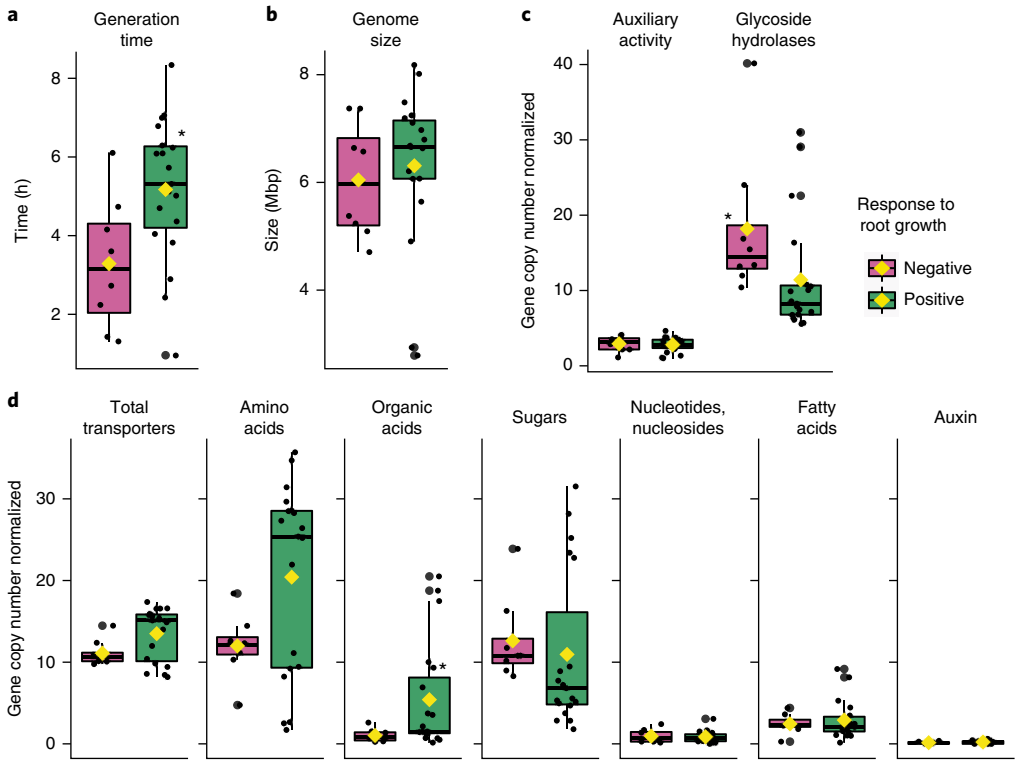

Fig. 3 | Distributions of select traits in the genomes of soil bacterial isolates classified into two groups based on the response to plant root growth. **a**, The minimum generation times predicted from genome sequences. **b**, Genome size of isolates. **c**, Extracellular enzymes for plant polymer degradation. **d**, Monomer transporters. All gene frequencies were adjusted for differences in genome size. The total number of transporters are shown as the percentage of transporters per genome. In each box plot, 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 the points excluding outliers. Outliers are 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 ($P < 0.05$; see Supplementary Data 2 for details) are denoted by *. Positive responders: $n = 19$; and negative responders: $n = 8$.

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 and 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 favoured 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 encoding organic acid transporters were significantly more abundant in positive responders than in negative 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 encoding glycoside hydrolases (GH) (primarily β-glucosidases (GH1, GH3 and GH5), β-xylosidases (GH43), β-glucanases (GH16), β-galactosidases (GH2), glucoamylases (GH15), α-glucosidases (GH13) and α-*N*-acetylgalactosaminidases (GH9) were more abundant in the genomes of negative responders (Fig. 3c and Supplementary Data 2) that are presumably better adapted to life outside the living root zone where easily accessible and assimilable substrates (for example, 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 suggest that traits related to this may be important to bacterial success in the rhizosphere. Thus, the chemical composition of plant exudates 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]; thus, we tested the composition of *A. barbata* exudates throughout its growth stages. For several reasons, we chose to use a hydroponic system to analyse the dynamics of *Avena* exudate chemistry. Using hydroponics is a trade-off[38], as 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 the 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

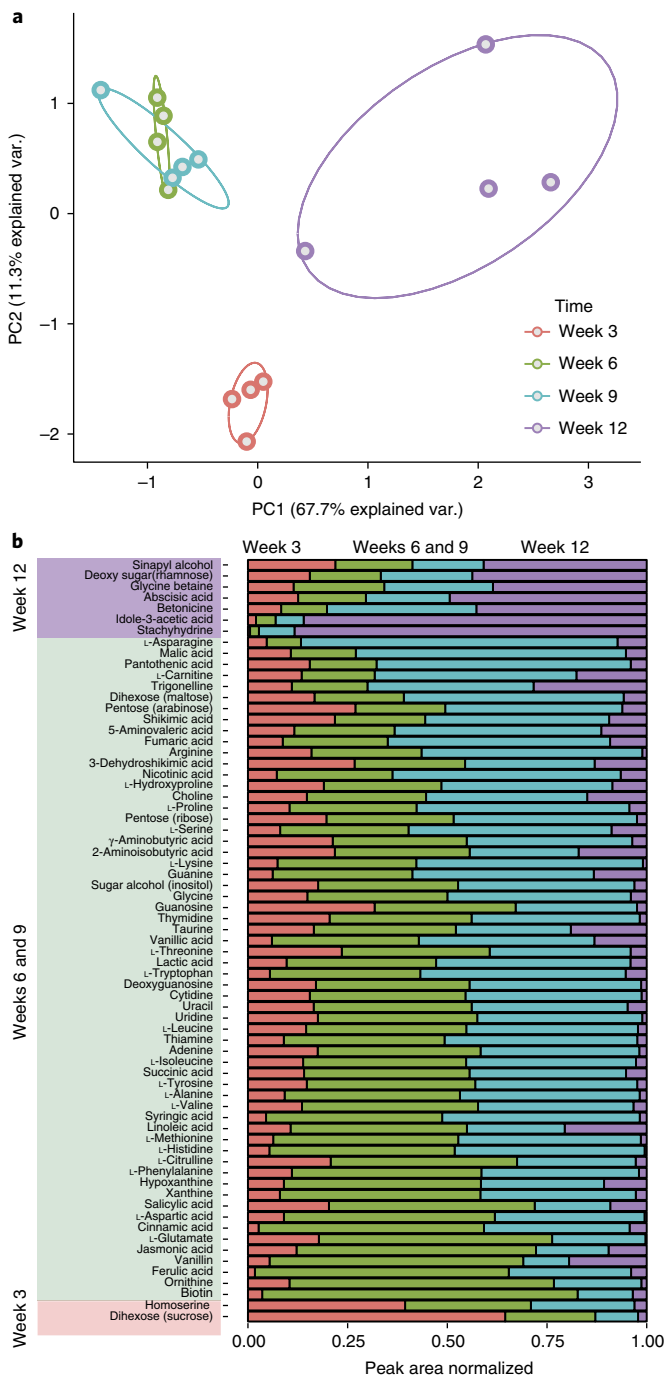

Fig. 4 | Changes in *A. barbata* exudation through plant development (weeks 3, 6, 9 and 12). a, Principle component (PC) analysis of the exudate profiles of *A. barbata* at each time point ($n = 16$ exudate profiles). **b**, Changes in the abundance of each exudate compound across plant developmental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

of *Avena* growth, which roughly proceeded along the same time scale as observed in soil. After collection, exudates were analysed 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 comprised a broad range of metabolites, including sugars, sugar alcohols, nucleotides, nucleosides, amino acids, organic acids, fatty acids, plant hormones and compatible solutes (Fig. 4 and Supplementary Data 3). Metabolite profiles from the early developmental stage (week 3) were distinct from those at week 6 and week 9, and from the late developmental stage corresponding to senescence (week 12) (Fig. 4a), whereas 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 and 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] and 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–microbial interactions and potential plant defence 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 and Supplementary Fig. 10). Plants release various aromatic compounds, as defence mechanisms against pathogens, in the form of signalling molecules and carbon sources for heterotrophs[44]. This increased release of aromatic compounds (vanillic acid, syringic acid, vanillin and 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 and 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-times 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 membranes, enzyme activity and to regulate detoxification of reactive oxygen species[45]. ABA has been suggested to have 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 various metabolites that change with root growth and plant developmental stages in a genetically programmed manner and suggest that changes in the 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 carbon content is reflected by individual compounds (undetected metabolites and larger water-soluble polymers remain unknown). Thus, although 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 carbon 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

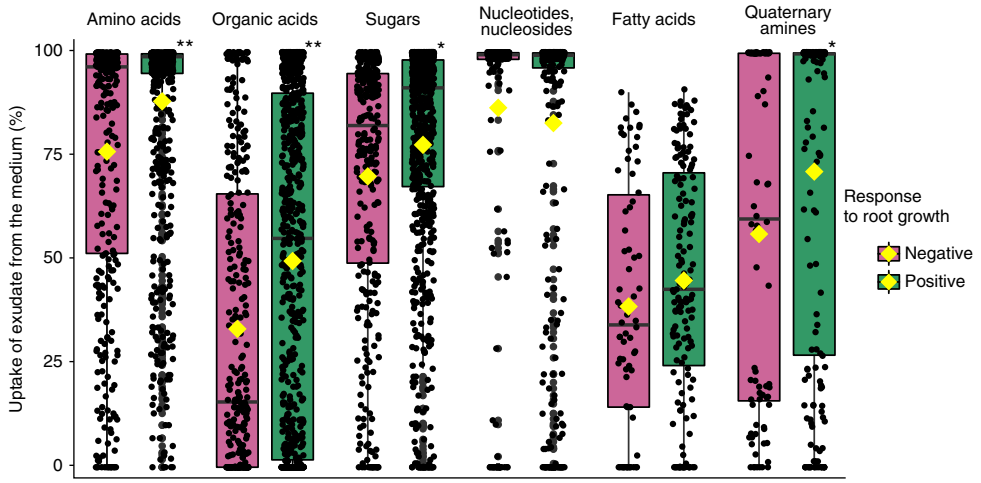

Fig. 5 | Distributions of root exudate metabolite uptake by isolates presented as the per cent uptake from the exudate medium. In each box plot, a point denotes a single metabolite and its per cent uptake in a single incubation. Diamond symbols in each box plot represent the mean. The box boundaries 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. Outliers are denoted as larger points outside whiskers. A significant difference between distributions were determined using the Kolmogorov–Smirnov test and are denoted with $^{**}P < 1 \times 10^{-8}$ and $^{*}P < 0.01$; see Supplementary Data 4 for details. Positive and negative responders to root growth: $n = 12$ isolates.

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 percentage 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 mixture, which is 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 an amino group), nucleotides and nucleosides, sugars, organic acids (organic acids that do not contain an amino group), quaternary amines and fatty acids. Isolates that were favoured 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 than negative responders.

These observations of enhanced amino acid 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 than bacteria that declined in response to root growth (Fig. 3d). This is compelling evidence that the selective uptake of organic acids from the mixture of exudates by rhizosphere bacteria interacts with the 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, the 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 the 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, the 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 that were 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 (the 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 that responded positively to the root had up to 48% higher percentage of metabolite depletion from the medium for the organic acids with aromatic rings than 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 the 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 *A. thaliana*[16], and together with γ-aminobutyric acid, salicylic acid concentration has been shown to correlate with specific taxa that are frequently enriched in the rhizosphere[14]. Taken together with our observations, it seems 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 and thymidine) were more preferentially consumed by the negative responders (Fig. 6a). Together, these data indicate that root exudate chemical composition selectively enriches rhizosphere responders

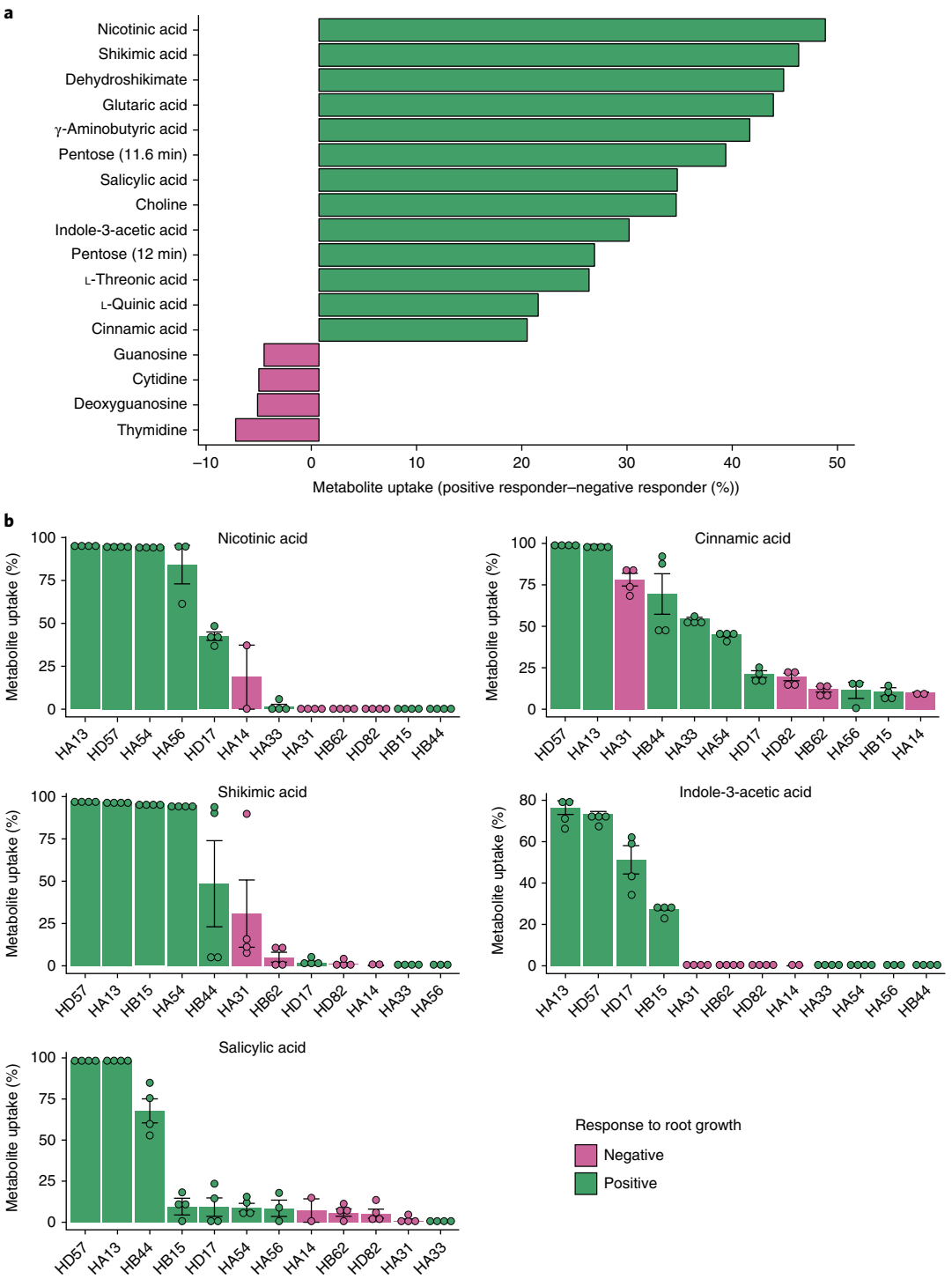

Fig. 6 | Substrate preferences of positive and negative responders. a, Metabolites with significant differences in their uptake from *Avena* exudates by isolates ($n=12$ isolates) ($P<0.05$, Kruskal–Wallis test; see Supplementary Data 4 for details). **b**, The uptake of aromatic organic acids from exudate media by isolates with positive ($n=8$ isolates) and negative ($n=4$ isolates) responses to plant growth. In each bar plot, isolates are shown on the *x* axis. A point denotes a biological replicate measurement of the per cent uptake of aromatic acid by an isolate. Metabolite uptake is given as a percentage of metabolite depletion by isolates from the medium, and error bars show the standard error of the mean. The key in panel **b** also applies to panel **a**.

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 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

Figs. 13b and 14a). Despite the limited number of observations used for the principal component regression, the predicted top metabolites that drive the 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 significant and these isolates were classified as 'undefined' (Fig. 2). We used substrate uptake preferences of these isolates as predictors to identify their behaviour in the 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 indicated by their 16S rRNA gene abundance patterns; this suggests that metabolite uptake traits may be particularly valuable predictors of rhizosphere colonization.

Conclusions

Interactions between roots, microorganisms and the soil matrix represent 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. *A. 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. *A. 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, *A. barbata*, to isolate associated soil bacteria and to measure plant exudation. To classify the dynamics of isolated bacteria, we used data on previously identified bacterial dynamics in response to *A. 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 paper, both of these species are referred to as *Avena*. Other soil properties have been previously described[2]. Growth media for the isolation of heterotrophic bacteria were formulated with different concentrations of nutrients, different solidifying agents and antioxidant enzymes (R2A 1/10, R2G 1/20, 1/100; OLI; VXylG) according to a previous study[57] (Supplementary Information). From 289 diverse isolates (Fig. 1), 39 were selected based on phylogeny and the relative abundance within the soil to be genome sequenced (Supplementary Figs. 1–3) using the Illumina HiSeq 2500 platform in accordance with the standard protocols of the US Department of Energy Joint Genome Institute (DOE JGI; Walnut Creek, CA, USA). Sequences of genomes were deposited, assembled and annotated at the JGI Integrated Microbial Genomes (IMG) portal (https://img.jgi.doe.gov; Supplementary Data 1 and 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 *A. fatua* have been previously reported[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 analysed 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 of $<1\times10^{-10}$ and ≥97% of gene sequence homology. Isolates that corresponded to OTUs detected in the soil were assigned the response pattern to plant growth of those OTUs, as previously determined[2].

Analysis of genomic features of bacterial isolates. We analysed 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 Figs. 3 and 4 and 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 a linear regression model (equation (1))[60].

$$\left(\Delta \mathrm{ENC}' = \frac{\mathrm{ENC}'_{\mathrm{all}} - \mathrm{ENC}'_{\mathrm{ribosomal\ protein\ genes}}}{\mathrm{ENC}'_{\mathrm{all}}}\right) \quad (1)$$

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

For extracellular enzyme traits, we calculated the gene copy numbers of glycoside hydrolases and auxiliary activity enzymes using a hidden Markov model search of protein sequences against the Carbohydrate Active enZYmes database (CAZy)[62,63]. For substrate uptake, transporters in the genomes were predicted using a hidden Markov model 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 (ref. [59]) and construction of a maximum likelihood tree using FastTree[65] (Supplementary Fig. 3).

***Avena* root exudate collection.** *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 6-litre hydroponic tubs with half-strength Murashige and Skoog[66] basal salt mixture M524 (Phyto Technology Laboratories). Hydroponic tubs were incubated at 24 °C on 16 h/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. 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 1 h in a growth chamber[67]. Milli-Q water containing exudates was immediately filter sterilized using 0.22-μm filters (Corning Inc.) and frozen. This exudate collection approach has been used previously for incubation times of up to 2.5 h 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 (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 per litre, magnesium sulfate 0.012 g per litre) 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, USA (0.025 mg carbon per gram dry soil) where these isolates originated (T. Whitman, personal communication). 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 per ml), this corresponds to approximately 125 mg carbon per litre. Root exudates were first lyophilized, redissolved in water and then added to growth media to achieve this final carbon concentration, as determined using a Shimadzu TOC-L Analyzer. 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 r.p.m. for up to 96 h. 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

concentrations of TOC in the uninoculated exudate medium (control samples) and TOC in spent exudate medium from isolate-inoculated treatments were quantified (Supplementary Fig. 6). Changes in the TOC concentrations in the spent media from the 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 per litre. Here, the necessary volume of cold (−20 °C) methanol containing internal standards (1 µg per ml 2-amino-3-bromo-5-methylbenzoic acid, 5 µg per ml 13C-15N-L-phenylalanine and 2 µg per ml 9anthracene carboxylic acid) was added to the dried exudates, and these were sonicated for 30 min using a ultrasonic bath (VWR). 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 6,000*g* for 1 min and filtered using 0.22-µm microcentrifuge PVDF filters to remove any particles and then analysed by LC-MS/MS.

MS analysis of exudates and spent bacterial growth media. Ultra-High Performance Liquid Chromatography (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). Full MS spectra were collected from *m/z* 70–1,050 at 70,000 FWHM (full-width at half-maximum) resolution, with MS/MS fragmentation data acquired using 10, 20 and 30 eV collision energies at 17,500 FWHM resolution. MS instrument parameters included sheath gas flow rate of 50 AU (arbitrary units), 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 × 2.1 mm, 5 µm, polymeric) at 40 °C at a flow rate of 0.25 ml min^{-1} with a 2 µl injection volume. The HILIC column was equilibrated with 100% buffer B (90:10 acetonitrile:H_2O with 5 mM ammonium acetate) for 1.5 min, diluting buffer B down to 50% with buffer A (H_2O with 5 mM ammonium acetate) for 23.5 min, down to 40% buffer B over 3.2 min, to 0% buffer B over 6.8 min, and followed by isocratic elution in 100% buffer A for 3 min. Exact mass and retention time coupled with MS/MS fragmentation spectra were used to identify compounds as described below.

Analysis of MS data. We first evaluated the overall quality of the data set by aligning the total ion chromatograms from all LC-MS runs using MZmine version 2.26 (ref. [68]) that demonstrated low variability across biological replicates (Supplementary Figs. 7 and 8). In addition, the 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 assessment (mass accuracy within 5 p.p.m., retention time and peak intensity of the internal standard), three samples that did not pass quality control were excluded from further analysis (Supplementary Fig. 9). Samples that passed the quality control steps were then analysed 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 of less than 5 p.p.m., retention time within 0.5 min and/or match of major MS/MS fragments). 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 (ref. [68]) was used for manual validation of identified metabolites (accurate mass of less than 5 p.p.m., 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 analysed under the same experimental conditions (for example, *m/z* and retention time; *m/z* and MS/MS; and 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 3 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 metabolite 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 the 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 an 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 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 and undefined indicating no significant change in relative abundance over time; Fig. 2 and 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- and 12-week-old plants) (Supplementary Fig. 10). Principal component analysis was used to evaluate and visualize the 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 was used to test for statistical differences in the distributions of uptake of specific chemical classes by the classified groups of bacterial isolates (that is, 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. Next, principal component analysis was performed and all principal components were used to build a principal component regression model based on training data 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 principal component regression model was selected by iterating through extracted PCs to obtain the lowest cross-validation error from 10-fold cross-validation (mean squared error of prediction, MSEP < 0.05)[76] (Supplementary Figs. 13a and 14a). 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 four 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].

Life Sciences Reporting Summary. Further information on experimental design is available in the Life Sciences Reporting Summary.

Data availability. The genomes of Hopland isolates are publicly available in the IMG (https://img.jgi.doe.gov) database under the IMG study name Mediterranean Grassland Soil Metagenome (MGSM): enabling a systems view of soil carbon and nitrogen biogeochemistry under a changing climate. The accession numbers of the genomes of isolates can be found in Supplementary Data 1. All data, including samples that we excluded from the analysis, were deposited to the Global Natural Products Social Molecular Networking (https://gnps.ucsd.edu) data repository (*Avena* exudates MSV000081804, bacterial isolate uptake of exudates MSV000081808 and root exudates components library MSV000081810).

Received: 3 May 2017; Accepted: 13 February 2018;
Published online: 19 March 2018

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 (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 (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 Biol. Biochem.* **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. USA* **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. *Environ. Microbiol.* **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. *J. Biol. Chem.* **288**, 4502–4512 (2013).
15. Shi, S. et al. Effects of selected root exudate components on soil bacterial communities. *FEMS Microbiol. Ecol.* **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 Environ.* **32**, 666–681 (2009).
19. Baetz, U. & Martinoia, E. Root exudates: the hidden part of plant defense. *Trends Plant Sci.* **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 Biol.* **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**, 509 (2013).
24. Dakora, F. D. & Phillips, D. A. in *Food Security in Nutrient-Stressed Environments: Exploiting Plants' Genetic Capabilities* (ed. Adu-Gyamfi, J. J.) 201–213 (Springer, Dordrecht, 2002). https://doi.org/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. *J. Exp. Bot.* **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 (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. USA* **110**, E1621–E1630 (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. *Curr. Opin. Biotechnol.* **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. *Environ. Microbiol. Rep.* **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**, 16160 (2016).
38. Neumann, G. & Römheld, V. in *The Rhizosphere* (eds Pinton, R. et al.) 23–72 (CRC Press, Boca Raton, 2009).
39. Walker, T. S., Bais, H. P., Grotewold, E. & Vivanco, J. M. Root exudation and rhizosphere biology. *Plant Physiol.* **132**, 44–51 (2003).
40. Salerno, G. Origin of sucrose metabolism in higher plants: when, how and why? *Trends Plant Sci.* **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. *Appl. Environ. Microbiol.* **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 Physiol.* **151**, 792–808 (2009).
43. Lemoine, R. et al. Source-to-sink transport of sugar and regulation by environmental factors. *Front. Plant Sci.* **4**, 272 (2013).
44. Lattanzio, V. & Lattanzio, V. in *Phytochemistry: Advances in Research* (ed. Imperato, F.) 23–67 (Research Signpost, Kerala, 2006).
45. Iqbal, N., Nazar, R. & Khan, N. A. (eds) *Osmolytes and Plants Acclimation to Changing Environment: Emerging Omics Technologies* (Springer, New York, 2015).
46. Noodén, L. D. (ed.) in *Senescence and Aging in Plants* 329–368 (Elsevier, Cambridge, MA, 1988). https://doi.org/10.1016/B978-0-12-520920-5.50016-X.
47. Brown, J. H., Paliyath, G. & Thompson, J. E. in *Plant Physiology* (ed. Steward, F. C.) 227–276 (Academic Press, San Diego, CA, 1991). https://doi.org/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 Physiol.* **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. *Environ. Microbiol.* **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. *Can. J. Microbiol.* **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. *Sci. Rep.* **5**, 13438 (2015).
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 Res.* **32**, W20–W25 (2004).
59. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **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 Res.* **42**, D490–D495 (2014).
63. Yin, C., Li, S. & Li, Q. Network traffic classification via HMM under the guidance of syntactic structure. *Comput. Netw.* **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 Res.* **35**, D274–D279 (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. *Physiol. Plant.* **15**, 473–497 (1962).
67. Khorassani, R. et al. Citramalic acid and salicylic acid in sugar beet root exudates solubilize soil phosphorus. *BMC Plant. Biol.* **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: a metabolite mass spectral database. *Ther. Drug. Monit.* **27**, 747–751 (2005).
71. Horai, H. et al. MassBank: a public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* **45**, 703–714 (2010).
72. Sumner, L. W. et al. Proposed minimum reporting standards for chemical analysis. *Metabolomics* **3**, 211–221 (2007).
73. *R: A Language and Environment for Statistical Computing* (R Development Core Team, 2017); www.R-project.org.
74. vegan: community ecology package. R package version 2.4-2 (Blanchet, F. G. et al., 2017); https://CRAN.R-project.org/package=vegan.
75. agricolae: statistical procedures for agricultural research. R package version 1.2-4 (de Mendiburu, F., 2016); https://CRAN.R-project.org/package=agricolae.
76. Mevik, B.-H. & Wehrens, R. The plsPackage: principal component and partial least squares regression in R. *J. Stat. Softw.* **18**, 1–23 (2007).

Acknowledgements

This study was supported by the DOE, Office of Science, Office of Biological Environmental Research, including a Genomic Sciences programme award no. DE-SC0010570, DE-SC0016247 and DE-SC0014079 to M.K.F. Work performed at the Lawrence Berkeley National Laboratory including DOE Early Career Awards to D.L. and T.R.N., and work performed at the DOE JGI (http://www.jgi.gov) and at the DOE Joint BioEnergy Institute (http://www.jbei.org) are supported by the DOE, Office of Science, Office of Biological and Environmental Research through Contract No. DE-AC02-05CH11231. D.L. was also supported in part by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 659910. Isolate genome sequencing was conducted by the DOE JGI, a DOE Office of Science User Facility, under a Community Science Program award to E.L.B., supported by the Office of Science of the DOE 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 B. Jargalsaikhan, R. Hossainkhil, D. Ly, S. Ouedraogo and Y. Nguyen for their assistance with maintenance of the bacterial isolate collection.

Author contributions

K.Z., T.R.N., M.K.F. and E.L.B. developed the hypotheses. K.Z., K.B.L., N.M., U.N.d.R., S.S. and D.L. performed the experimental analyses. K.Z., H.C., U.K., Z.H., U.N.d.R. and B.P.B. analysed the data. K.Z., T.R.N., M.K.F. and E.L.B. wrote the paper. All authors provided comments and edits on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41564-018-0129-3.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to T.R.N. or E.L.B.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Brodie, Northen

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

For 16S rRNA we had 12 biological replicates for each of five time points (week 0, 3, 6, 9, and 12). This sample size was selected based on the optimal sample replication number for the 16S rRNA community analysis in our previous study. We used 39 genome-sequenced isolates that represent abundant and phylogenetically representative part of the microbial community of the rhizosphere of Avena to analyze genomic traits of the rhizosphere bacteria.

To identify substrate uptake preferences of the rhizosphere isolates we analyzed exometabolome of 16 isolates with four biological replicates. This number of replicates is the minimum number of replicates required for the statistical significance testing.

To identify dynamics of plant exudates we used four biological replicates collected at four time points that correspond main developmental stages of Avena (3, 6, 9, and 12). This number of replicates is the minimum number of replicates required for the statistical significance testing.

Sample size and the selection criteria are reflected in Methods section and in the Supplementary Figure 1.

2. Data exclusions

Describe any data exclusions.

We excluded samples from the exometabolome analysis by using three criteria:
1. Isolate had poor bacterial growth based on the Optical Density of the sample.
2. Isolate did not consume carbon from the medium based on the Total organic carbon content in inoculated medium compared to the control medium.
3. Sample did not pass LC-MS quality control based on the intensity of the internal standard.

3. Replication

Describe whether the experimental findings were reliably reproduced.

For 16S rRNA we had 12 biological replicates for each of five time points (week 0, 3, 6, 9, and 12).

For exometabolomic analysis of exudates we used four biological replicates collected at four time points ((week 3, 6, 9, and 12).

For exometabolomic analysis of isolates we used four biological replicates for each isolate.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Isolates and genome-sequenced isolates were classified into two main groups based on the change in their abundance - positive responders and negative responders.

For the LC-MS analysis samples were randomized for the injection sequence on the LCMS plate.

After LC-MS analysis samples were classified by using statistical approach into two main groups - positive responders and negative responders.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

During isolation of the bacteria for the experiment we were blinded to any group allocation of these isolates.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

For the genomic analysis we used BLAST, MUSCLE, HMM search and FastTree.
For metabolite analysis we used MZmine and Metabolite Atlas package (https://github.com/biorack/metatlas).
For statistical analysis we used R.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

Isolate genome sequences are available at the JGI IMG portal (https://img.jgi.doe.gov)

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

n/a

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

n/a

b. Describe the method of cell line authentication used.

n/a

c. Report whether the cell lines were tested for mycoplasma contamination.

n/a

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

n/a

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

n/a

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

n/a