Title
Niche differentiation is spatially and temporally regulated in the rhizosphere

Journal
The ISME Journal: Multidisciplinary Journal of Microbial Ecology, 14(4)

ISSN
1751-7362

Authors
Nuccio, Erin E
Starr, Evan
Karaoz, Ulas
et al.

Publication Date
2020-04-01

DOI
10.1038/s41396-019-0582-x

Peer reviewed

Niche differentiation is spatially and temporally regulated in the rhizosphere

Erin E. Nuccio[1], Evan Starr[2], Ulas Karaoz[3], Eoin L. Brodie[3], Jizhong Zhou[3,4,5], Susannah Tringe[6], Rex R. Malmstrom[6], Tanja Woyke[6], Jillian F. Banfield[3,7], Mary K. Firestone[3,7], and Jennifer Pett-Ridge[1]

[1]Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore CA, 94551 USA; [2]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720 USA; [3]Earth and Environmental Sciences Area, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA; [4]Institute for Environmental Genomics, Department of Botany and Microbiology, University of Oklahoma, Norman, OK 73019 USA; [5]State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing, 100084, China; [6]DOE Joint Genome Institute, Walnut Creek, CA 94598, USA; [7]Department of Environmental Science, Policy and Management, University of California, Berkeley, CA, 94720 USA

Running title: Decomposition gene expression in the rhizosphere

Corresponding authors: Erin Nuccio nuccio1@llnl.gov 925-423-9983, Jennifer Pett-Ridge pettridge2@llnl.gov 925-424-2882, Lawrence Livermore National Laboratory, Livermore CA, 94551 USA

Abstract

The rhizosphere is a hotspot for microbial C transformations, and the origin of root polysaccharides and polymeric carbohydrates that are important precursors to soil organic matter. However, the ecological mechanisms that underpin rhizosphere carbohydrate depolymerization are poorly understood. Using *Avena fatua*, a common annual grass, we analyzed time-resolved metatranscriptomes to compare microbial function in rhizosphere, detritusphere, and combined rhizosphere-detritusphere habitats. Population transcripts were binned with a unique reference database generated from soil isolate and single amplified genomes, metagenomes, and stable isotope probing metagenomes. While soil habitat significantly affected both community composition and overall gene expression, succession of microbial functions occurred at a faster time scale than compositional changes. Using hierarchical clustering of upregulated decomposition gene expression, we identified four distinct microbial guilds populated by taxa whose functional succession patterns suggest specialization for substrates provided by fresh growing roots, decaying root detritus, the combination of live and decaying root biomass, or aging root material. Carbohydrate depolymerization genes were consistently upregulated in the rhizosphere, and both taxonomic and functional diversity were high in the combined rhizosphere-detritusphere—suggesting coexistence of rhizosphere guilds is facilitated by niche differentiation. Metatranscriptome-defined guilds provide a framework to model rhizosphere succession and its consequences for soil carbon cycling.

INTRODUCTION

The rhizosphere is a critical zone for C transformations in the terrestrial biosphere, since roots are the primary source of soil organic matter (1-3) and can significantly alter the rate of soil C turnover (4-6). Plants deposit a significant proportion of their photosynthates into soil as root biomass or exudates (7), and plant-derived polymeric carbohydrates such as cellulose and hemicellulose are the most abundant polysaccharides in soil (8, 9). These rhizodeposits create a high resource, high activity environment, and stimulate a bloom of microbial biomass (10) that undergoes ecological succession as roots grow and senesce (11, 12), selecting for organisms that benefit mineral nutrition (13) and overall plant health (14). Rhizodeposits also stimulate depolymerization by cellulases, chitinases, and proteases (15, 16) leading to higher rates of decomposition, and thus nutrient availability, in the region surrounding both living roots and decaying root detritus (17). However, the ecological controls of rhizosphere carbohydrate depolymerization are not well understood, which limits our ability to accurately model soil C dynamics (18) and plant-microbe interactions.

Previous studies suggest that rhizosphere community assembly is due to selective processes such as niche differentiation or habitat filtering (19-21), and metagenome sequences indicate the rhizosphere selects for microbial genomes with functional capacities that are distinct from bulk soil, where carbohydrate active enzyme (CAZy) genes are enriched in the rhizosphere (19, 21-23). However, it is unclear if this large genomic potential translates to high carbohydrate degradation activity in the environment. Genomic composition represents the full functional repertoire of a microorganism, the "fundamental metabolic niche" that constrains all the potential habitats it could hypothetically occupy (24, 25). But microbial communities contain functional redundancy that is not necessarily realized or expressed in the ecosystem (26). To understand "realized" metabolic niches within complex rapidly changing microbial communities (26), it is essential to consider expressed functional measurements—such as transcripts, proteins, or metabolites—that can reflect niche differentiation in real time (27, 28).

Measurements of expressed functions provide a useful way to study community assembly based on shared activities rather than shared phylogeny, and allow us to define microbial guilds—cohorts of

64 organisms defined by similar function that is not dependent on phylogeny (29). In cases like the
65 rhizosphere and detritusphere, where communities might logically be defined by functional traits rather
66 than taxonomic relatedness, guilds defined by gene expression, rather than species, may be the most
67 relevant parameter for understanding patterns of diversity (30), modeling community interactions (31),
68 and identifying the gene transcripts that mediate root-accelerated decomposition. However, the ideal
69 parameters for identifying or operationally defining guilds in microbial communities are unresolved.
70 Microbial guilds have been identified previously by adding single substrates to soil and measuring
71 subsequent increases in taxonomic relative abundance (32). We theorize that functional guilds can also be
72 identified using population-resolved gene expression, where guild members turn on and off genes in a
73 coherent spatial or temporal manner in response to the same habitat, resources, or environmental
74 perturbations. Genome-centric analyses now allow us to track transcription in individual populations,
75 which may be a more relevant approach than grouping transcripts across disparate classes or phyla (33).
76 Using comparative metatranscriptomics, we studied microbial degradation of macromolecular plant
77 compounds, hypothesizing that gene expression would reflect distinct functional succession patterns in
78 different soil habitats (rhizosphere and detritusphere), consistent with niche differentiation. The
79 transcripts were extracted from soil near live and decaying roots in microcosms containing *Avena fatua*, a
80 common annual grass, growing in its native soil. Population transcripts measured over the course of three
81 weeks were binned using a genome-resolved reference database specific to our experimental soil. We
82 found that carbohydrate depolymerization was executed by a series of microbial guilds, with distinct
83 spatial and temporal response patterns in gene expression. We tested whether these guilds had differing
84 life history traits based on their preferred substrate (rhizosphere or detritusphere), and assessed whether
85 carbohydrate depolymerization expression was controlled by: a) increasing population size, b)
86 upregulating transcription, or c) synergistically upregulating transcription in response to combined
87 resources (i.e., combined rhizosphere-detritusphere). Our work provides a mechanistic framework for
88 understanding the drivers of rhizosphere succession and identifies carbohydrate/lignolytic gene transcripts
89 mediating root-accelerated decomposition.
90

METHODS

Experimental Design

The annual grass, wild oat (*Avena fatua)* was grown in two-chamber microcosms with a sidecar region designed to allow access to the rhizosphere (Fig. S1) (10, 16, 34). The outer wall of the sidecar was clear plastic, allowing us to monitor root growth and rhizosphere age. Microcosms were packed with soil (1.2 g/cm^3) collected beneath a stand of *A. barbata* at the Hopland Research and Extension Center (Hopland, CA, USA). The soil is a Bearwallow-Hellman loam, pH 5.6 with 2% total C (35). For half the microcosms, 50 g soil was amended with 0.4 g dried *A. fatua* root detritus and spread on top of 100 g of sidecar soil; root detritus (also called 'root litter') had been grown in sand, triple washed, aged 1 year, and chopped to 1 mm. Each microcosm also contained a 1 µm mesh 'bulk soil' bag, designed to exclude roots but allow moisture equilibration; these contained 2 g soil, either amended with 0.016 g detritus (bulk + detritus) or unamended (bulk). Plants were grown in the main chamber for six weeks before starting the experiment. Six days prior, the divider separating the main chamber and sidecar was replaced with a slotted divider, and microcosms were tilted 40°, allowing roots to grow into the sidecar.

Root age was tracked in order to collect rhizosphere soil of defined age. New root growth was marked three days after the microcosms were tilted, and harvests took place 3, 6, 12 or 22 days later (Fig. S1). At each timepoint, we destructively sampled paired rhizosphere and bulk soil for two treatments (with and without detritus) with 3 biological replicates, collecting 48 total samples (24 rhizosphere, 24 bulk).

Sample Collection

Rhizosphere soil <2 mm from the root was excised with a scalpel. Root sections and adhering soil were placed immediately in ice cold Lifeguard Soil Preservation Reagent (MoBio), vortexed for 2 min on medium speed, and pelleted according to the Lifeguard protocol. Roots were removed using flame-

sterilized tweezers and supernatant removed. Pelleted soils were frozen on dry ice and stored at -80°C. Bulk soils were processed identically. Approximately 1 g of soil was collected per sample.

The remaining sidecar soil was collected for edaphic characterization. Soil pH was measured as per Fierer and Jackson (36) with a Corning 340 pH meter. Gravimetric moisture content was determined by measuring water loss from 10 g fresh soil after 2 days at 105°C. Total carbon (TC) was measured on a subset of the samples to calculate C addition due to the root material using an elemental analyzer IRMS (PDZ Europa, Limited, Crewe, UK).

DNA/RNA Extraction

DNA and RNA were co-extracted from 0.5 g of frozen soil using a phenol-chloroform extraction protocol (37, 38). DNA and RNA were separated using the Qiagen AllPrep kit. RNA was treated with TURBO DNase (Thermo Fisher Scientific) following the manufacturer's protocol and concentrated by ethanol precipitation. RNA was visualized an Experion Electrophoresis System (Bio-Rad), and quantified using the Qubit RNA BR Assay Kit (Thermo Fisher Scientific).

Single Amplified Genomes

Rhizosphere and root endophyte microbial cells were sorted to create single amplified genomes (SAGs). Three grams of roots coated in rhizosphere soil were washed in 10ml of cell release buffer (0.5% Tween, 2.24 mM Na pyrophosphate in PBS), vortexed for 1 min at full speed on a horizontal shaker, and soft pelleted at 2KxG for 2 min (repeated 4 times). Cells from the supernatant were stained with SYBR, imaged using a Zeiss Axioimager M2 (CNR Biological Imaging Facility, UC Berkeley), and counted using ImageJ (39). Glycerol was added to the supernatant at a final concentration of 15% and frozen at –80° C prior to sorting. Roots were washed in PBS, shaken in 10% bleach (1 min), rinsed in water, dried by centrifugation, and frozen at –80° C prior to maceration using a sterile mortar and pestle prior to cell sorting. Thawed rhizosphere and root endophyte cells were isolated using fluorescence activated cell

sorting, amplified by multiple displacement amplification, and screened using 16S rRNA sequencing (40).

Sequencing Library Preparation

Metatranscriptomes, iTags (16S, ITS), and single amplified genomes (SAGs) were sequenced at the Joint Genome Institute (JGI); see Supplemental Methods for full details. Briefly, for metatranscriptomic libraries, ribosomal RNA was depleted using the Ribo-Zero rRNA Removal Kit (Epicentre) for Plants and Bacteria and reverse transcribed into cDNA. cDNA was sequenced (2x150bp) on an Illumina HiSeq2000 sequencer using a TruSeq SBS sequencing kit (v3). For iTag analysis, paired DNA and RNA were amplified from the same nucleic acid extract prepared for metatranscriptomics. iTag libraries targeted the bacterial 16S V4 region (primers 515F, 805R) (41, 42) and the fungal ITS2 region (primers ITS9, ITS4) (43, 44) using barcoded reverse primers (41). Amplicons were sequenced (2x300bp) on an Illumina MiSeq sequencing platform using a MiSeq Reagent Kit (v3 600 cycle). Selected SAGs that successfully amplified 16S rRNA were sequenced using the Illumina NextSeq platform (40).

Sequence Processing

Metatranscriptomic raw reads were quality-trimmed (Q20) using fastqTrimmer, and artifacts were removed using DUK (45). Contaminating ribosomal RNA and transfer RNA were identified and removed with bowtie2 (46) by mapping reads against SILVA (47), Greengenes (48), IMG rRNA, GtRNAdb (49), and tRNADB-CE (50) databases. In total we sequenced 408 Gbp of RNA, and after *in silico* contaminant filtering, we obtained an average of 43 million paired-end metatranscriptomic reads per library (see Table S1 for repository IDs and sequencing statistics). We did not detect any bias towards rhizosphere or bulk soil in either sequencing library size (Figure S2) or gene diversity (Figure S3).

Amplicons were analyzed on JGI's iTag analysis pipeline (iTagger) (41), which created OTUs at the 97% and 95% identity level for bacterial 16S and fungal ITS, respectively. Contaminants were removed using

DUK, merged with FLASH (51), and dereplicated. Dereplicated sequences were sorted by decreasing abundance, clustered with USEARCH (52), and assigned taxonomy using the RDP classifier (53). SAG sequences were processed using BBTools (54). Sequences were filtered using BBDuk and mapped against masked contaminant references (human, cat, dog) using BBMap and BBMask. Reads with an average kmer depth <2 were removed. Normalization was performed with BBNorm and error correction with Tadpole (54). Sequences were assembled using SPAdes (version v3.7.1) (55); 200bp was trimmed from contig ends; contigs were discarded if length was < 2kbp or read coverage was < 2. Cross-contaminated contigs were identified and removed using CrossBlock (54). Automated SAG decontamination was performed with ProDeGe (version 2.3) (56) and assemblies were discarded if the total size was < 200 kbp. Details about the final draft assemblies are listed in Table S2.

Soil-Specific Reference Database

Transcripts were mapped against a genome database specific to our Hopland CA experimental soil, composed of 96 metagenome-assembled genomes (MAGs) (NCBI PRJNA517182), 221 MAGs from a stable isotope probing (SIP) rhizosphere density gradient (57) (http://ggkbase.berkeley.edu/), 39 isolate genomes (58), and 29 single amplified genomes (SAGs) (this study; Table S1). Reference genomes were dereplicated using whole pairwise genome alignments at 98% nucleotide identity (59); we selected the highest quality representative based on completeness of single copy genes (60). To ensure we did not include multiple fragmented genomes from the same organism, genomes > 70% complete were clustered into groups that overlapped by at least 50%; genomes < 70% complete were clustered in a second round using a 30% overlap. The highest quality representative was selected for each cluster (score = # single copy genes – 2 * multiple single copy genes; the genome with the highest N50 was selected to break a tie). This resulted in 335 total genomes for our custom reference database (Table S3), composed of 214 rhizosphere SIP-MAGs (64%), 53 soil MAGs (16%), 39 isolate genomes (12%), and 29 SAGs (9%).

Gene Annotation and Counts

Gene prediction was performed on all genome bins using Prodigal in metagenome mode (61). Protein sequences were annotated using dbCAN2 (62) (accessed April 2017), KEGG (63), and ggKbase (http://ggkbase.berkeley.edu/). Proteins with CAZyme functional domains were manually curated to generate a consensus annotation: CAZymes without KEGG or ggKbase annotations were ignored, and if the KEGG and ggKbase annotations disagreed, KEGG was selected. Genes containing signal peptide signatures for extracellular protein transport were annotated using SignalP 4.1 (64).

Trimmed and filtered reads were mapped against our soil-specific database using BBSplit (54). Reads that mapped ambiguously to multiple reference genomes were discarded to prevent double-counts. Transcripts were binned into population transcriptomes using a relaxed similarity cutoff (80% min identity) and should not be interpreted as genome transcriptomes. Gene counts were determined using featureCounts (R package: Rsubread).

Data Analysis

Metatranscriptomic and amplicon sequencing data were normalized using DESeq2 to account for differences in sequencing effort (65), except for Shannon diversity analysis, where reads were rarified and diversity indices calculated using QIIME 1.9.1 (66). At each time point, significant differential expression relative to bulk soil was determined using DESeq2, which adjusts p values for multiple comparisons. Ordination and graph visualization were conducted in R (67). Data were ordinated using non-metric multidimensional scaling (R package: vegan), and significantly different clusters were determined using adonis (68). Correlations between environmental data and ordination data were tested using envfit (R package: vegan).

Carbohydrate Depolymerization CAZymes (d-CAZy)

We selected population transcriptomes with 4+ upregulated carbohydrate depolymerization genes for further analysis, which identified 26 of the 335 populations. Target substrates for depolymerization

CAZymes (d-CAZy) were initially classified based on Berlemont and Martiny (69) and then refined into the following putative substrate categories using the consensus annotation described above: cellulose, xylan, xyloglucan, pectin, plant polysaccharides, microbial cell walls, starch and glycogen, xylose and cellulose oligosaccharides, oligosaccharides, mono- and disaccharides (see Table S7 for gene names and references). Area-proportional Venn diagrams were created to visualize the subset of the community that significantly upregulated d-CAZy transcripts relative to the total d-CAZy genomic potential (R package: venneuler).

Guild Assignment

We defined guilds based on d-CAZy expression over time and across treatments. Average d-CAZy differential gene expression (log2 fold change) relative to bulk soil was visualized using heatmaps (R package: pheatmap). One-dimensional hierarchical clustering was used to assign heatmap groups, which were classified as guilds based on resource preference and timing of peak gene expression. Maximum gene expression per enzyme per treatment was plotted by reference genome using ggplot2. Phylogeny is presented in accordance with current taxonomic nomenclature (70).

Decomposition Strategies

Decomposition strategies were assessed by evaluating gene expression levels relative to population abundance, and by comparing expression in the rhizosphere or detritusphere to the combined rhizosphere-detritusphere. Expression levels of the housekeeping genes Gyrase A and B were used as proxies for population abundance, as gyrases often have stable expression patterns over a variety of treatment conditions (71). Populations that increased in abundance, with significantly higher gyrase expression relative to bulk soil (by DESeq2), where increased d-CAZy expression is partially attributable to a larger population size were assigned to the ‘Grower’ strategy. Populations where d-CAZy expression was upregulated above population size (based on per capita gene expression), and d-CAZy fold-change was 3-fold higher than gyrase fold-change relative to bulk soil were assigned to the ‘Upregulator’ strategy. We

use 'Synergist' to describe populations with 3-fold higher gene expression when combined resources were available (i.e., combined rhizosphere-detritusphere) compared to the rhizosphere or detritusphere alone.

RESULTS

Microcosm soil properties

Root detritus additions increased soil carbon from 2.0% ± 0.1 to 2.8% ± 0.1. Gravimetric soil moisture at the time of harvest averaged 0.34 ± 0.067 g water g^{-1} dry soil, with the exception of the final timepoint, when microcosms had an average of 0.11 ± 0.013 g water g^{-1} dry soil due to plant transpiration (Figure S4a). The addition of root detritus significantly increased bulk soil pH by 0.14 pH units at the first timepoint, and this differenced decreased over time (Figure S4b).

Rapid community and functional assembly in the rhizosphere and detritusphere

Living roots and detritus rapidly altered bacterial community structure and functional assembly. All treatments diverged from bulk soil within 3 days, as seen by clear groupings in NMDS ordination space for both 16S cDNA and mRNA transcripts (Figure 1a and 1c, respectively). For 16S cDNA, the rhizosphere and detritusphere significantly shaped community composition (Figure 1) (see Table S4 for PERMANOVA F tables). Changes due to time explained 19% of the community variability, indicating that some taxonomic succession occurred within the treatments (Figure S5a). In contrast, fungal community composition, measured by ITS cDNA, was indistinguishable between rhizosphere and bulk soil ($p > 0.1$), and was instead significantly altered by both the detritus amendment and time (Figure 1b) (Table S4).

Time was the dominant factor structuring bacterial gene expression; transcripts from the final (22-day) timepoint clearly separate from earlier time points for all treatments (Figure 1c) (Table S4). This shift is

correlated with both soil moisture (envfit: $r^2 = 0.87$, $p < 0.001$) (Figure S5a) and time (envfit: $r^2 = 0.57$, $p < 0.001$) (Figure S5b), likely because evapotransporation due to increased root biomass caused soil drying across all treatments at the final timepoint. When only the first three time points are considered (days 3, 6, 12), the rhizosphere and detritus treatments are the dominant factors structuring community gene expression (Table S4).

Early colonists of the rhizosphere rapidly increased in relative abundance within 3 days (16S cDNA, Table S5), and included Proteobacteria (Burkholderiaceae) and Verrucomicrobiota (Opitutaceae). Early colonists of the detritusphere included Fibrobacterota, Verrucomicrobiota (Chthoniobacteraceae, Opitutaceae), Armatimonadota, Bacteroidota, and Proteobacteria. In contrast, relatively few Actinobacteria and Acidobacteria significantly responded to either the rhizosphere or detritusphere on this timescale (3-22 days).

Root detritus increased taxonomic and functional diversity

Root detritus amendment increased the taxonomic and functional Shannon diversity of both rhizosphere and bulk soil, and the combined rhizosphere-detritusphere had the highest overall taxonomic diversity by the final timepoint (Tukey HSD analysis, Figure 2a). Shannon taxonomic diversity was calculated based on 16S rRNA genes. Root detritus amendment of bulk soil significantly increased KEGG functional diversity, and appeared to have a similar effect on rhizosphere soil, although the trend was not significant at $p < 0.05$ (Tukey HSD analysis, Figure 2b). Of our four treatments, rhizosphere soils (with and without root detritus) had the highest expressed functional diversity after 22 days of root growth.

Roots stimulated expression of carbohydrate depolymerization transcripts

We curated a set of CAZyme genes relevant for plant and microbial carbohydrate depolymerization (d-CAZy, Table S7) and assessed expression of carbohydrate depolymerization transcripts relative to the bulk soil treatment. Overall, rhizosphere communities had the most significantly upregulated d-CAZy

296 genes (Figure 2c). The combined rhizosphere-detritusphere had the largest number of significantly
297 upregulated d-CAZy genes, with the exception of the final time point, when unamended rhizosphere soil
298 had the largest number of upregulated genes. This result was generally consistent across the four major
299 CAZyme classes (auxiliary activity, carbohydrate esterases, glycoside hydrolases, polysaccharide lyases)
300 (Figure S6). In the bulk soil, root detritus additions initially stimulated a large pulse of d-CAZy activity
301 but this dropped dramatically over time; by the final timepoint, only 10-20% of the genes were
302 distinguishable from bulk soil expression levels.
303
304 **Realized niches in rhizosphere and detritusphere**
305 The fundamental metabolic niche describes the full metabolic repertoire of a microorganism, and is
306 represented by its total genomic content. For the individual populations in our reference database, we
307 identified populations with statistically significant gene expression and compared it to genomic content in
308 an effort to identify the 'realized' metabolic niches within our bacterial community (Figure 3). While
309 many populations had the genomic capacity for carbohydrate depolymerization, only a small fraction
310 significantly upregulated these genes relative to bulk soil by DESeq2. Populations upregulating
311 cellulases and xylanases relative to bulk soil across the three treatments were 11-15% and 10-19% of the
312 total genomic potential, respectively (Figure 3). The relative expression patterns for gyrase A and B
313 housekeeping genes indicate that the general population dynamics followed similar patterns by treatment
314 as observed for functional genes.
315
316 **Guilds defined by temporal and habitat gene expression**
317 We assessed population gene transcription patterns over time and across habitats to define ecological
318 guilds. The majority of differential gene expression came from 26 of our 335 reference genomes; these
319 had at least four d-CAZy genes significantly upregulated relative to bulk soil. We note that 20 of them
320 were derived from our rhizosphere SIP-metagenome database (Figure 4). These bacterial populations had
321 distinct rhizosphere versus detritusphere transcriptional preferences (one dimensional hierarchical

clustering, Figure 4). We averaged d-CAZy differential expression per population to show broad differential expression patterns (Figure S7). Using statistical upregulation of depolymerization genes as a proxy for resource preference and guild membership, we defined "Rhizosphere," "Detritusphere," and "Aging Root" guilds, and a "Low Response" group where there was no discernable habitat preference.

Carbohydrate depolymerization guilds undergo functional succession

Transcriptionally-defined guilds captured a functional succession in carbohydrate depolymerization, for both polysaccharides and also oligosaccharide breakdown products. The Rhizosphere and Detritusphere guilds had high d-CAZy expression within the first 6 days, then between 12-22 days an additional Aging Root guild emerged (Figure 4).

The Rhizosphere guild contained Proteobacteria order Burkholderiales and a Verrucomicrobiota population from the Opitutaceae (Figure 4, Group 1). Cellulases (endoglucanases), xylanases, and xyloglucanases were most highly expressed at 3 days, as were enzymes for potential breakdown products like cellulose- and xylan-oligosaccharide hydrolases (beta-glucosidases and beta-xylosidases, respectively) (Figure S7a-d). One Burkholderiaceae population did not follow this pattern, and instead had high d-CAZy expression at the final timepoint. Overall, xyloglucan hydrolases were characteristic of rhizosphere populations, and observed only once in the detritusphere (bulk + detritus) (Figure 5).

The Detritusphere guild was phylogenetically diverse, including members from Proteobacteria, Myxococcota, Fibrobacterota, Bacteroidota, and Actinobacteroita phyla (Figure 4, Group 2). With the exception of the Rhizobiaceae population, members of the Detritusphere guild typically upregulated cellulases and xylanase (or both) soon after detritus was added (3, 6 days), and cellulose- or xylan-oligosaccharide hydrolases for potential breakdown products (Figure 5).

In the Aging Root guild, Actinobacteriota populations from the *Streptomycetaceae* and Catenulisporales had high d-CAZy gene expression at the final timepoint (22 days) in the rhizosphere, and early gene expression in the detritus-amended treatments (Figure 4, Group 3). The Aging Root guild had almost no upregulated genes for starch, glycogen, cell wall, and disaccharide decomposition (Figure 5).

Guild-based assessment of decomposition strategies

We used metatranscriptomic expression patterns to determine the prevalence of three decomposition strategies within the microbial guilds (Figure 4): a.) increased population abundance, b.) upregulated gene expression (above per capita abundance), or c.) synergistic gene upregulation when combined resources were available (i.e., combined rhizosphere-detritusphere). We interpreted significant gyrase upregulation relative to bulk soil as a population increase (DESeq2, Figure S8).

Decomposition strategies varied by guild membership, and were not mutually exclusive. All members of the Rhizosphere and Detritusphere guilds were "Growers" and increased in abundance (Figures 4 and S8). Rhizosphere guild populations were also "Upregulators" (3-fold higher gene expression per capita) (Figures 4 and S9). In contrast, enzyme expression in the Detritusphere guild tracked population levels; only one population was an "Upregulator". Over half of the Aging Root guild did not change in abundance relative to bulk soil. This guild, composed entirely of Actinobacteriota and two *Steptomyces* populations, had "Upregulators" with populations sizes statistically indistinguishable from bulk soil.

Three "Synergist" populations upregulated gene expression in response to combined resources, with 3-fold higher gene expression in the combined rhizosphere-detritusphere compared to either habitat alone (Figures 4 and S10). These include Verrucomicrobiota (Opitutaceae) and Burkholderiales populations within the Rhizosphere guild, and a Fibrobacterota population from the Detritusphere guild. The 30S ribosomal protein S3 (RP-S3) from the Opitutaceae reference genome was 93% similar to the *Opitutus terrae* RP-S3 (by blastx (72)), an obligate anaerobe isolated from rice paddy soil. The Burkholderiales

and Fibrobacterota reference genomes were most closely related to uncultivated MAGs (RP-S3 85% similar to Rhizobacter MAG and 73% similar to Fibrobacterota MAG, respectively). Both the Optitutaceae and Fibrobacterota reference genomes contained CBB3 cytochrome oxidases (putatively, a microaerophilic version of cytochrome oxidase); these genes were actively expressed but not upregulated relative to bulk soil. At the early timepoints, the Opitutaceae population upregulated enzymes for xylan degradation (arabinoxylan arabinofuranosidase) and xylan breakdown products (xylan 1,4-beta xylosidase, alpha-D-xyloside xylohydrolase) (Figures 5 and S7d). The Fibrobacterota population synergistically upregulated endoglucanases, endo-1,4-beta-xylanases, and enzymes targeting their potential breakdown products (beta-xylosidase, cellobiose phosphorylase) (Figure S7h). At later timepoints, the Burkholderiales upregulated putative lignocellulosic enzymes such as endoglucanase, tannase and feruloyl esterase, and rhamnogalacturonan lyase (Figure S7e).

DISCUSSION

Rapid community and functional assembly in the nascent rhizosphere and detritusphere

The soil microbial community surrounding roots undergoes a compositional succession corresponding to the phenological stages of plant growth (11, 12). However little is known about microbial gene expression during rhizosphere succession, and the temporal relationship between functional succession versus community changes. We used genome-centric, comparative metatranscriptomics to identify taxa mediating root-enhanced decomposition using carbohydrate gene transcripts. Since many soil taxa are non-cultivable by conventional methods, this approach offers insights into the physiologies of populations only known by sequencing (33). Our results, some of the first using a genome-centric metatranscriptome approach in soil, illustrate that different microbial populations have specialized functions and life strategies based on spatiotemporal differences in root habitats.

We found community and functional assembly proceeded at different rates—while taxonomic composition underwent minor successional changes over three weeks, expressed functional composition

399 distinctly shifted between 12-22 days. mRNA has a short half-life relative to DNA, and is a sensitive
400 indicator about ongoing ecological processes and near-real-time conditions experienced by cells (73). Our
401 previous work indicates rhizosphere community composition continues to shift from three weeks until
402 senescence (11), but the faster changes we observe for transcript inventories suggest microbes experience
403 changes in rhizodeposits, environmental conditions (e.g. moisture, pH, O_2), or other signals on the scale
404 of days (77). The relative speed of functional shifts suggests that expressed functional succession occurs
405 at a faster time scale than compositional changes, and possibly dictates the rhizosphere microbial
406 community succession that occurs over longer time scales. This illustrates a benefit of using time-
407 resolved metatranscriptomics to assess activity of specific microbial populations and the processes that
408 lead to community assembly, since organisms transcriptionally respond to stimuli on a shorter time scale
409 than evinced by replication.
410
411 *SIP-metagenomes produced the most useful genomes for soil metatranscriptomics*
412 The proportion of population-specific transcripts mapping to our reference genomes illustrates the
413 comparative benefits of the four sequence products in our custom database: single amplified genomes,
414 isolate genomes, deeply sequenced bulk soil metagenomes, and rhizosphere stable isotope probing (SIP)
415 metagenomes. Genomes derived from rhizosphere SIP-metagenomes proved to be the most relevant for
416 transcript mapping, and the source of most of the populations with 4+ upregulated carbohydrate
417 depolymerization genes. In a previous study (57), these populations also showed high ^{13}C-incorporation
418 by rhizosphere ^{13}C-SIP (unpublished data), where the plants were continuously-labeled with $^{13}CO_2$ for 6
419 weeks. Our results indicate that SIP-metagenome datasets may be a highly fruitful genomic resource for
420 environmental metatranscriptomics and other omics analyses.
421
422 *Metatranscriptomic guilds provide a framework to understand rhizosphere succession*
423 By assigning expressed carbohydrate depolymerization genes to individual population genomes derived
424 from our custom genome database, we stepped beyond gene-centric studies that have shown rhizosphere

gene expression with plant development (12) or environmental changes (74) and identified specific carbohydrate depolymerization guilds based on shared spatiotemporal gene expression.

We used these guilds to evaluate decomposition strategies that underpin altered carbohydrate degradation rates commonly found surrounding roots (4). In macroecology, the guild concept is a common way to group populations as functional ecological units, based on their resource utilization traits or life history strategies (25, 29). In environmental microbiology, while next-generation sequencing allows us to group microbial communities taxonomically (75), taxonomy and function may not correspond (26, 30, 76). We identified four guilds based on spatiotemporal CAZy expression patterns (Rhizosphere, Detritusphere, Aging Root and Low Response). This ecological categorization framework may be particularly useful for phylogenetically ubiquitous microbial functions—for example, soil organic matter decomposition (32, 76-78), stress, or nitrogen mineralization—where guilds are based on shared life history traits rather than phylogeny.

Within each guild, many populations engaged multiple catabolic pathways for carbohydrate degradation, including potential degradation by-products such as cellulose- and xylose-oligosaccharides. For example, populations from the Rhizosphere and Detritusphere guilds not only expressed enzymes for cellulose and xylan degradation, but also their breakdown products. Recent work suggests that facilitative processes such as cross-feeding in large networks can act to stabilize coexisting competitors for resources (79-81). Genome-resolved metagenomic analyses indicate the importance of metabolic byproduct handoff in linking together interacting members of microbial communities (80, 81). The breadth of carbohydrate degradation pathways that Rhizosphere and Detritusphere guilds engage in may be a potential explanation for the stable, positive and repeatable interaction networks in the rhizosphere that we observed in a previous study (82). We hypothesize that complex cross-feedings networks promote coexistence within highly interconnected rhizosphere communities (79).

Niche differentiation promotes coexistence of rhizosphere and detritusphere guilds

By combining taxonomy and function, genome-resolved metatranscriptomics, we examined both the 'potential' and 'realized' metabolic niches (26, 33) of bacteria in our experiment. The niche differentiation concept asserts that organisms coexist by subdividing available resources, such as food or space (25, 32). As the number of niches increases in a system, so should the number of coexisting species (83). During root colonization in the combined rhizosphere-detritusphere, coexisting guilds began to develop, demonstrating that carbohydrate depolymerization preferences that were evident when root habitats were presented in isolation (rhizosphere or detritusphere), could coexist when combined. When detritus was added to the rhizosphere, the most populations predominantly demonstrated spatial and temporal coexistence rather than synergistic consumption of resources. These results are reflected by the higher functional alpha diversity in the combined rhizosphere-detritusphere; we saw approximately additive increases in functional diversity when a new resource (root detritus) was added to the system. Our work suggests that spatial and temporal niche differentiation promotes microbial coexistence in the rhizosphere and detritusphere.

Guild-based assessment of carbohydrate degradation strategies

We further assessed if our guilds had differing d-CAZy transcriptional strategies and evaluated if our increases in gene expression (a) tracked increases in population size ('Growers'), (b) were upregulated per capita ('Upregulators'), or (c) synergistically upregulated when both root exudates and detritus were available ('Synergists'). These strategies were not mutually exclusive, and their prevalence varied according to guild membership. All organisms in the Rhizosphere guild were both 'Growers' and 'Upregulators,' while the Detritusphere guild were primarily 'Growers.' Multiple studies have shown that the input of organic-C substrates can increase or decrease the rates of C degradation of surrounding soil organic matter, which is a phenomenon known as priming (4, 5, 84). Due to the large number of significantly upregulated decomposition transcripts in the rhizosphere, both with and without detritus amendments, this system has a high potential for increased rates of decomposition in the rhizosphere, as

was previously observed in this plant-soil system (6). This is consistent with the expectations for positive rhizosphere priming, where fresh organic matter provided by the rhizosphere stimulates the production of enzymes that can degrade soil organic matter (17, 85, 86).

Based on the rhizosphere priming hypothesis, we would expect to observe "Synergists" in the combined rhizosphere-detritusphere. Strikingly, we only observed 3 "Synergist" populations, and two of these were putative microaerophiles. This suggests that these populations may also be partitioning their niches based on changes in the edaphic environment, such as oxygen or pH, rather than simply by consuming combined resources alone. The Verrucomicrobiota MAG is distantly related to Opitutaceae isolates derived from oxygen-limited rice patties and insect guts (87-90). Fibrobacteres include cellulose degrading bacteria found in mammal rumens (91), termite guts (92), anaerobic cellulose reactors (93), and rice paddy soil (94). Both MAGs contain cytochrome oxidases with high oxygen affinity (CBB3), which is associated with organisms living in microaerophilic environments (95). High amounts of heterotrophic respiration can create microaerophilic zones in otherwise aerobic environments, such as the rhizosphere (96, 97). The combined oxygen demand from both the rhizosphere and detritusphere may have been sufficiently high to create microaerophilic niche for root detritus decomposition, thus providing a possible mechanism for the observed synergistic response. Both of these populations are rhizosphere inhabitants found in our other studies (11, 20) suggesting that this synergistic decomposition in the combined rhizosphere-detritusphere may be functionally significant in semiarid grasslands.

During the functional succession of guilds, one guild emerged during the latter half of the experiment as the rhizosphere aged. Interestingly, more than half of the Aging Root guild had population sizes that were indistinguishable from bulk soil based on gyrase housekeeping gene expression, but in some cases still were "Upregulators." Similarly, by 16S analysis, few Actinobacterial taxa changed in relative abundance in response to the treatments. This suggests that these populations were actively utilizing carbohydrates and not appreciably changing their population sizes over the timescale we measured. A

503 recent SIP study on forest soils found that Actinobacteriota only accumulated ^{13}C after 21 days (98).
504 Since SIP requires replication to increase enrichment of DNA (99), the authors hypothesized that this
505 could be due to slow growth. Our results support this hypothesis, and indicate that populations with
506 minimal growth can still be active and functionally relevant in the community (100). We also note that
507 some Actinobacterial populations in the Detritusphere guild had significant d-CAZy transcription as early
508 as 3 days. Thus metatranscriptomics provides a way to assess functional relevance that is independent of
509 changes in taxonomic relative abundance.
510
511
512 **CONCLUSIONS**
513 Niche differentiation is central to theories of coexistence (25, 32, 83). Recent advances in metagenomic
514 sequencing have allowed us to define the fundamental metabolic niches of representatives from poorly-
515 known phyla, for whom there are little phenotypic data (57, 101). Using genome-centric
516 metatranscriptomics to define the realized metabolic niches for soil populations, we found that
517 carbohydrate depolymerization guilds rapidly emerged during rhizosphere community assembly. Using
518 these guilds, we determined the prevalence of three d-CAZy transcriptional strategies in the rhizosphere,
519 and found that rhizosphere organisms upregulate decomposition transcripts in addition to increasing
520 population sizes. Further, these populations used both primary and breakdown products, and supports
521 recent observations that metabolic handoffs link together interacting members of microbial communities
522 (79-81). Guilds dynamics of carbohydrate depolymerization during rhizosphere succession provides a
523 key step towards developing microbially-constrained models to predict the fate of soil carbon.
524
525 **ACKNOWLEDGEMENTS**
526 This research was supported by the U.S. Department of Energy Office of Science, Office of Biological
527 and Environmental Research Genomic Science program under Awards SCW1589 and SCW1039 to JPR,
528 DE-SC0010570 and DOE-SC0016247 to MKF, and DOE-SC10010566 to JFB. Sequencing was

conducted as part of Community Sequencing Awards 1487 to JPR and 1472 to MKF. Work conducted at Lawrence Livermore National Laboratory was supported under the auspices of the U.S. Department of Energy under Contract DE-AC52-07NA27344. Work conducted at Lawrence Berkeley National Laboratory and the U.S. DOE. Joint Genome Institute, a DOE Office of Science User Facility, was supported under Contract No. DE-AC02-05CH11231. Cell imaging was conducted at the CNR Biological Imaging Facility at UC Berkeley. We thank Shengjing Shi, Katerina Estera, Jia Tian for assisting with sample processing and Alex Probst for providing bioinformatics support and advice.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

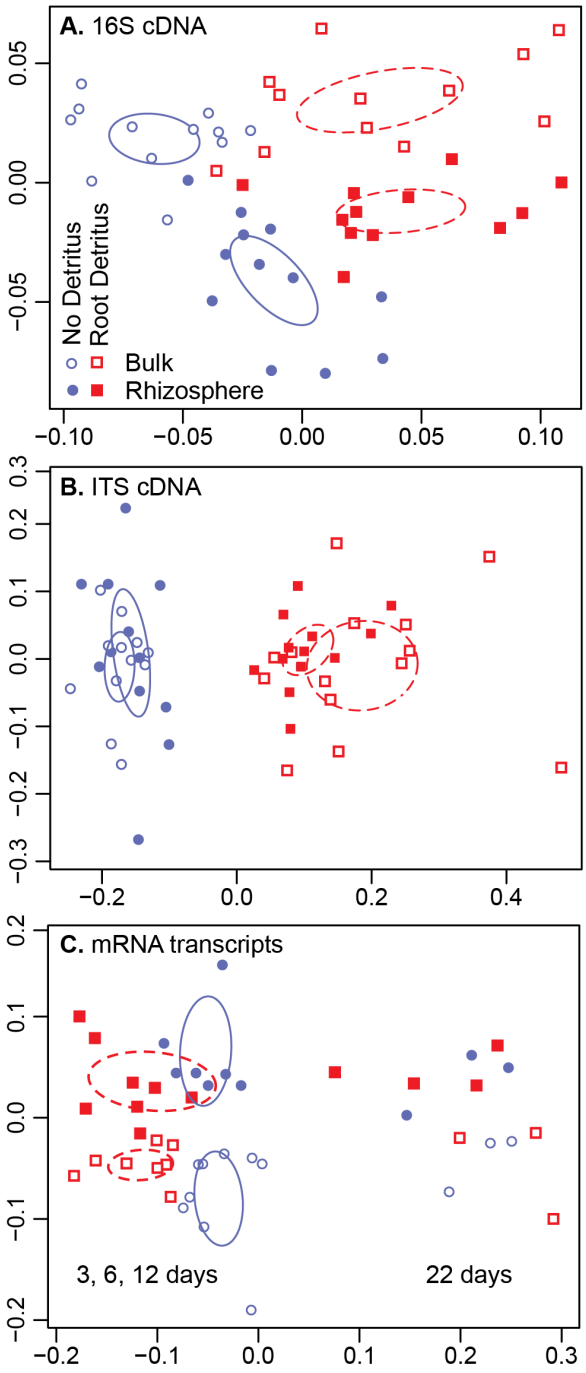


Figure 1. Influence of living roots and root litter on soil microbial communities and their gene expression during 3 weeks of *Avena fatua* root growth (independent harvests at 3, 6, 12, 22 days), as represented by NMDS ordination. Microbial community composition was measured by (A) bacterial 16S cDNA amplicons and (B) fungal ITS cDNA amplicons. Expressed functional composition was measured by (C) mRNA transcripts. Symbols represent four experimental habitats: rhizosphere (filled symbols), bulk soil (hollow symbols); each with added root detritus (red), or without added root detritus (blue). Ellipses represent the standard error of the weighted average of the centroid scores (calculated by ordiellipse). n=3 for each habitat and timepoint.

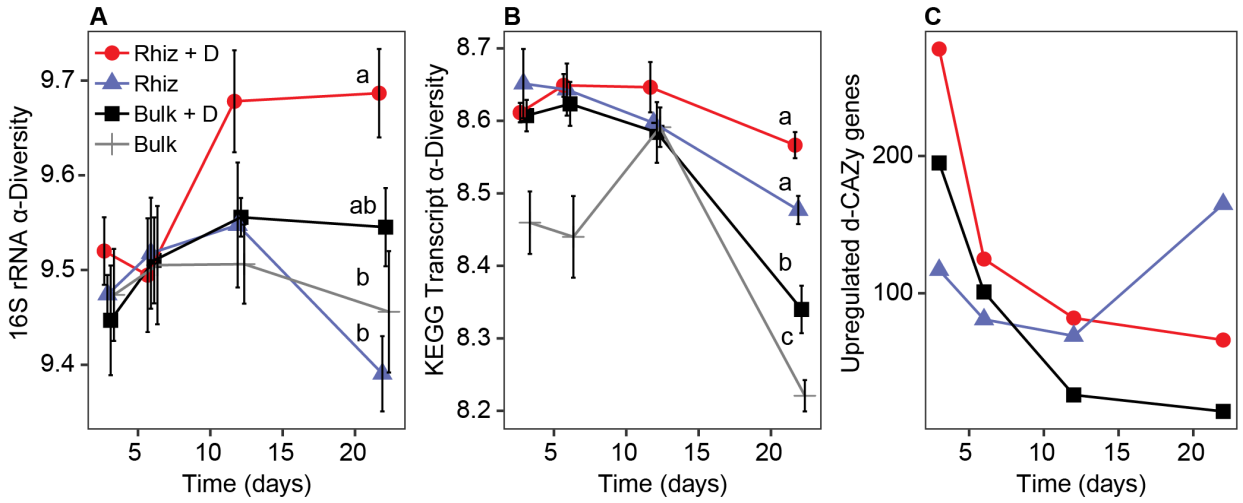


Figure 2. Taxonomic versus functional diversity in rhizosphere and bulk soils with and without detritus (+ D) harvested from *Avena fatua* microcosms over the course of 22 days. Average Shannon diversity for (A) 16S rRNA genes, and (B) KEGG functional genes derived from community mRNA transcripts. Error bars reflect one standard error. In order to make our results more comparable to a prior study of bacterial succession in the *Avena* rhizosphere, Shannon taxonomic diversity was calculated based on 16S rRNA genes (11). The different letters represent significant differences measured by Tukey HSD analysis at the final timepoint. (C) The cumulative number of significant differentially-upregulated decomposition CAZy (d-CAZy) genes relative to bulk soil, measured by DESeq. Treatments are: rhizosphere + detritus (red circle), rhizosphere (blue triangle), bulk soil + detritus (black square), and untreated bulk soil (grey cross).

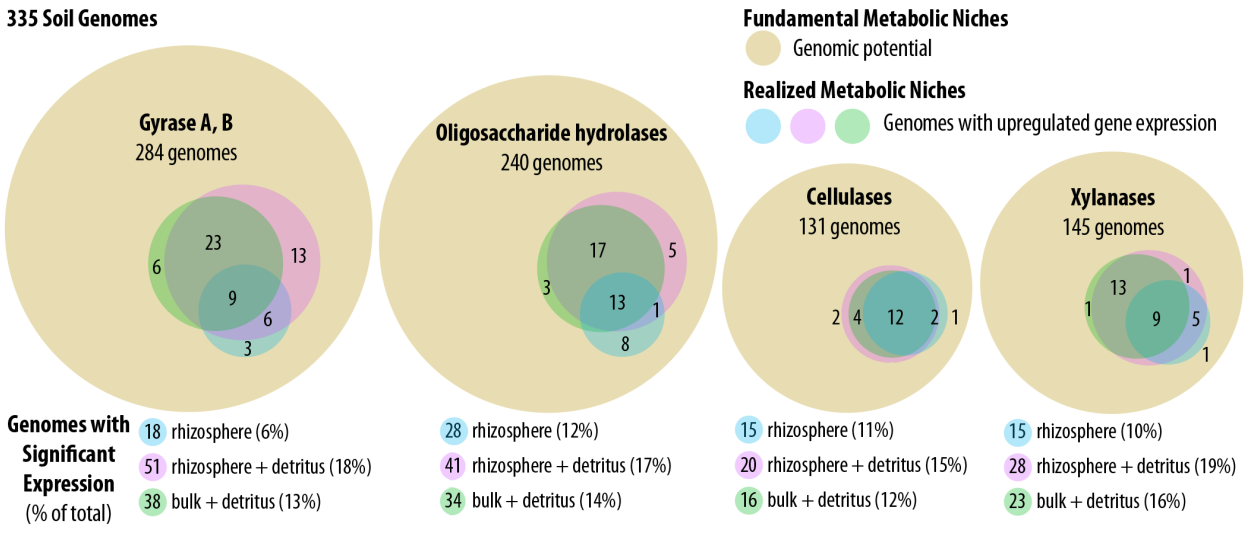


Figure 3. Fundamental (genomic potential) versus realized metabolic niches (upregulated gene expression) for key carbohydrate degradation gene classes. Area-proportional Venn diagrams indicate the number of functionally active taxa by soil habitat relative to the total metagenomic capacity for 335 assembled soil genomes. The outer circle (brown) indicates the number of unique genomes in the reference database with the genomic potential for the specified class of genes; inner circles reflect the number of taxa that differentially upregulated each class of genes relative to bulk soil for each treatment: rhizosphere (blue), rhizosphere + detritus (pink), and bulk soil + detritus (green). Overlapping regions represent shared niche space, with the number of genomes shared between different treatments. Genome classes analyzed include: gyrase A, B (housekeeping gene), oligosaccharide hydrolases (e.g., glycosidases, xylanases), cellulases, and xylanases (see Table S7 for full gene list). The bottom panel lists the number of active genomes by treatment, and the percentage of active genomes relative to total genomic potential is denoted in parentheses.

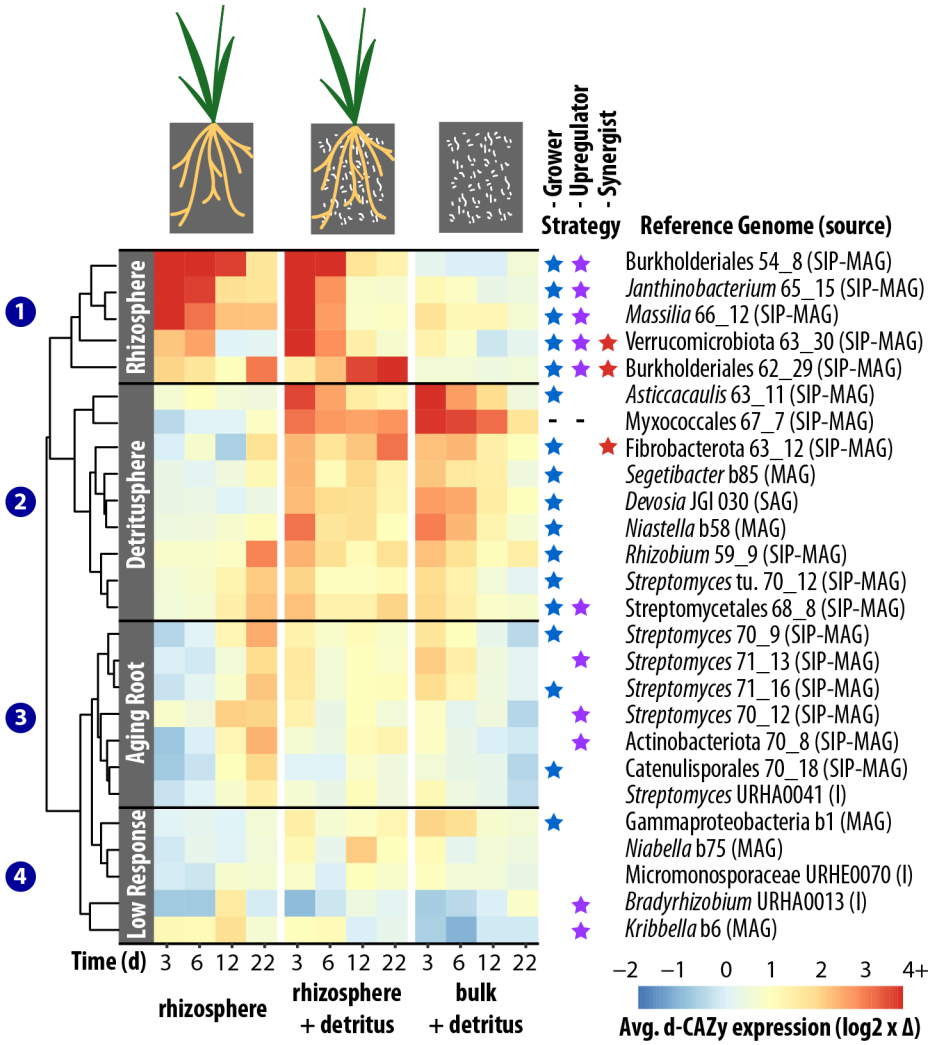
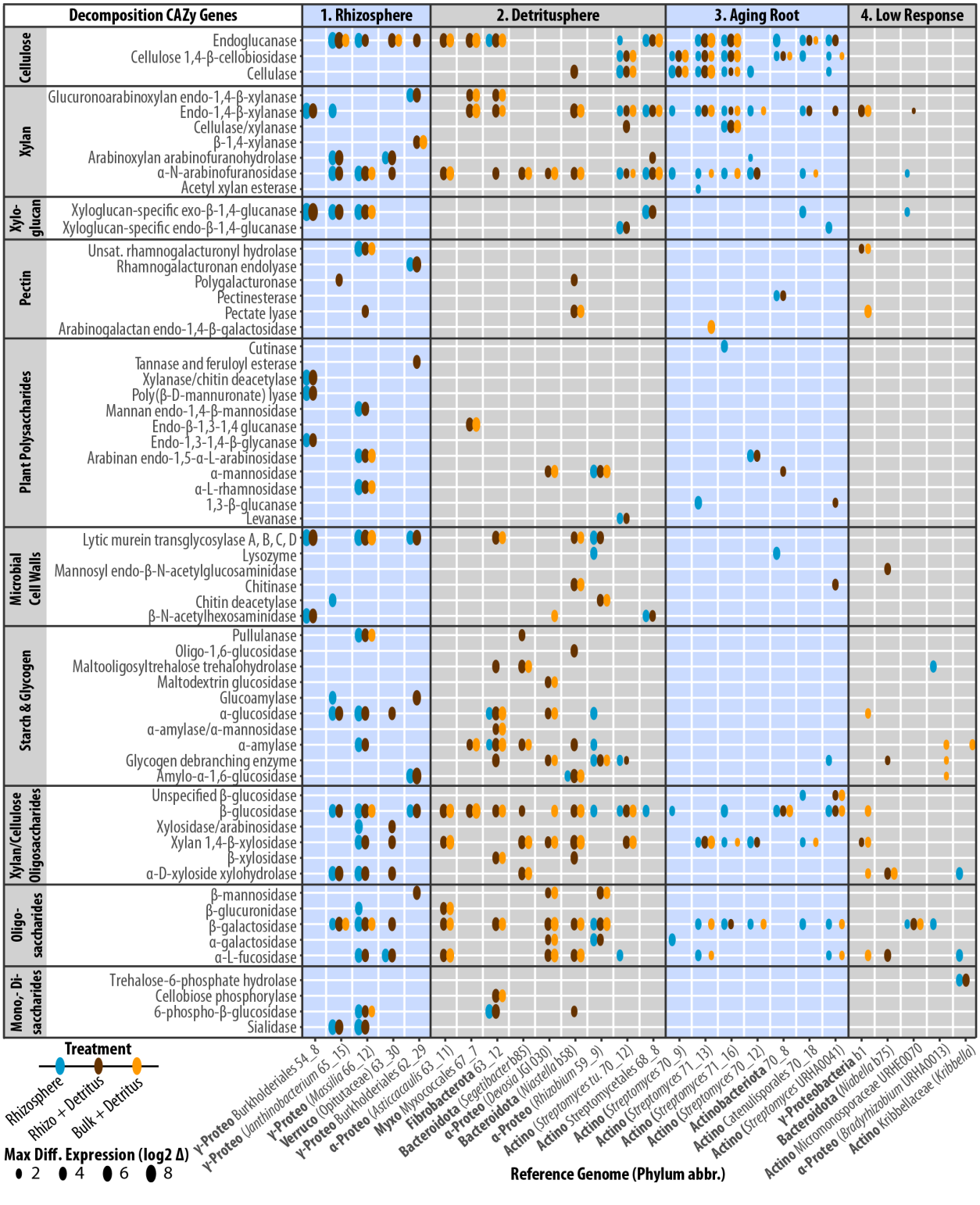


Figure 4. Time-series heatmap representing average decomposition CAZy (d-CAZy) gene expression per genome for 26 d-CAZy-responsive taxa during a 22-day *Avena fatua* microcosm experiment; responsive taxa significantly upregulate 4+ d-CAZy genes relative to bulk soil. Red indicates log2-fold gene upregulation in the treatment, blue indicates gene upregulation in bulk soil. Reference genome taxonomy is listed for the population transcriptomes (rows), as is the source of the genome: rhizosphere SIP-metagenome (SIP-MG), soil metagenome (MG), cultured isolate genome (I), single amplified genome (SAG). Time (days) is indicated by the columns. Metatranscriptomic guild assignment was accomplished through one-dimensional hierarchical clustering and is denoted by the left gray bars and numbers; high d-CAZy gene expression when living roots were present were assigned to the 'Rhizosphere' guild; high d-CAZy expression when added detritus was present formed the 'Detritusphere' guild; high d-CAZy expression when living roots were present, but where expression peaked at the last timepoint, formed the 'Aging Root' guild. Stars indicate decomposition strategy: blue stars indicate populations that significantly increased abundance ('Growers,' Fig. S8)); purple stars indicate populations where per capita gene expression was 3x > abundance ('Upregulators,' Fig. S9)); red stars indicate 'synergist' populations, where gene expression in combined rhizosphere-detritusphere was 3x > than rhizosphere or detritusphere alone ('Synergist,' Fig. S10)). Hyphens indicate no gyrase data was available for the calculation.



607 **Figure 5**. Upregulated decomposition CAZy genes for 26 bacteria classified into decomposition guilds
608 defined in this study (Rhizosphere, Detritusphere, Aging Root, Low Response; see Fig. 4). Ovals and
609 their size indicate maximum differential expression relative to bulk soil (log2-fold change) over the time
610 course for the treatments: rhizosphere (blue ovals), rhizosphere + detritus (brown ovals), and bulk +
611 detritus (yellow ovals). Genes are grouped by the enzyme's putative target substrate: plant
612 polysaccharides (cellulose, xylan, xyloglucan, pectin, other plant polysaccharides), microbial cell walls,
613 starch and glycogen, xylan- and/or cellulose-oligosaccharides, other oligosaccharides, and mono- and di-
614 saccharides. Phylum abbreviations: γ-Proteobacteria (γ-Proteo), Verrucomicrobiota (Verruco), α-
615 Proteobacteria (α-Proteo), Myxococcota (Myxo), Actinobacteriota (Actino).
616
617

618 **References**

619 1. Rasse DP, Rumpel C, Dignac M. Is soil carbon mostly root carbon? Mechanisms for a specific
620 stabilisation. *Plant Soil* 2005; **269**: 341-56.
621 2. Pett-Ridge J, Firestone MK. Using stable isotopes to explore root-microbe-mineral interactions in
622 soil. *Rhizosphere* 2017; **3**: 244-53.
623 3. Sokol NW, Bradford MA. Microbial formation of stable soil carbon is more efficient from
624 belowground than aboveground input. *Nature Geoscience* 2019; **12**: 46-53.
625 4. Cheng W, Parton WJ, Gonzalez Meler MA, Phillips R, Asao S, McNickle GG, et al. Synthesis
626 and modeling perspectives of rhizosphere priming. *New Phytol* 2014; **201**: 31-44.
627 5. Huo C, Luo Y, Cheng W. Rhizosphere priming effect: A meta-analysis. *Fungal Genet Biol* 2017;
628 **111**: 78-84.
629 6. Bird JA, Herman DJ, Firestone MK. Rhizosphere priming of soil organic matter by bacterial
630 groups in a grassland soil. *Soil Biol Biochem* 2011; **43**: 718-25.
631 7. Pausch J, Kuzyakov Y. Carbon input by roots into the soil: Quantification of rhizodeposition
632 from root to ecosystem scale. *Global Change Biol* 2018; **24**: 1-12.
633 8. Eichorst SA, Kuske CR. Identification of cellulose-responsive bacterial and fungal communities
634 in geographically and edaphically different soils by using stable isotope probing. *Appl Environ Microbiol*
635 2012; **78**: 2316-27.
636 9. Berlemont R, Martiny AC. Phylogenetic distribution of potential cellulases in bacteria. *Appl*
637 *Environ Microbiol* 2013; **79**: 1545-54.
638 10. DeAngelis KM, Brodie EL, DeSantis T, Andersen G, Lindow S, Firestone MK. Selective
639 progressive response of soil microbial community to wild oat. *ISME J* 2009; **3**: 168-78.
640 11. Shi SJ, Nuccio E, Herman DJ, Rijkers R, Estera K, Li JB, et al. Successional trajectories of
641 rhizosphere bacterial communities over consecutive seasons. *Mbio* 2015; **6**: e00746-15.
642 12. Chaparro JM, Badri DV, Vivanco JM. Rhizosphere microbiome assemblage is affected by plant
643 development. *ISME J* 2014; **8**: 790-803.

644 13. van der Heijden MGA, Bardgett RD, van Straalen NM. The unseen majority: soil microbes as
645 drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol Lett* 2008; **11**: 296-310.
646 14. Chaparro JM, Sheflin AM, Manter DK, Vivanco JM. Manipulating the soil microbiome to
647 increase soil health and plant fertility. *Biol Fertility Soils* 2012; **48**: 489-99.
648 15. Spohn M, Kuzyakov Y. Spatial and temporal dynamics of hotspots of enzyme activity in soil as
649 affected by living and dead roots-a soil zymography analysis. *Plant Soil* 2014; **379**: 67-77.
650 16. DeAngelis KM, Lindow S, Firestone MK. Bacterial quorum sensing and nitrogen cycling in
651 rhizosphere soil. *FEMS Micobiol Ecol* 2008; **66**: 197-207.
652 17. Kuzyakov Y, Blagodatskaya E. Microbial hotspots and hot moments in soil: Concept & review.
653 *Soil Biol Biochem* 2015; **83**: 184-99.
654 18. Wang K, Peng C, Zhu Q, Zhou X, Wang M, Zhang K, et al. Modeling global soil carbon and soil
655 microbial carbon by integrating microbial processes into the ecosystem process model TRIPLEX-GHG. *J*
656 *Adv Model Earth Sy* 2017; **9**: 2368-84.
657 19. Mendes LW, Kuramae EE, Navarrete AA, van Veen JA, Tsai SM. Taxonomical and functional
658 microbial community selection in soybean rhizosphere. *ISME J* 2014; **8**: 1577-87.
659 20. Nuccio EE, Anderson-Furgeson J, Estera K, Pett-Ridge J, de Valpine P, Brodie EL, et al. Climate
660 and edaphic controllers influence rhizosphere community assembly for a wild annual grass. *Ecology*
661 2016; **97**: 1307-18.
662 21. Ofek-Lalzar M, Sela N, Goldman-Voronov M, Green SJ, Hadar Y, Minz D. Niche and host-
663 associated functional signatures of the root surface microbiome. *Nat Commun* 2014; **5**: 4950.
664 22. Levy A, Salas Gonzalez I, Mittelviefhaus M, Clingenpeel S, Herrera Paredes S, Miao J, et al.
665 Genomic features of bacterial adaptation to plants. *Nat Genet* 2018; **50**: 138-50.
666 23. Yan Y, Kuramae EE, de Hollander M, Klinkhamer PGL, van Veen JA. Functional traits dominate
667 the diversity-related selection of bacterial communities in the rhizosphere. *ISME J* 2017; **11**: 56-66.
668 24. Hutchinson GE. Concluding Remarks. *Cold Spring Harb Symp Quant Biol* 1957; **22**: 415-27.
669 25. Silvertown J. Plant coexistence and the niche. *Trends Ecol Evol* 2004; **19**: 605-11.

670 26. Louca S, Polz MF, Mazel F, Albright MBN, Huber JA, O'Connor MI, et al. Function and
671 functional redundancy in microbial systems. *Nat Ecol Evol* 2018; **2**: 936-43.
672 27. Tannock GW, Wilson CM, Loach D, Cook GM, Eason J, O'Toole PW, et al. Resource
673 partitioning in relation to cohabitation of *Lactobacillus* species in the mouse forestomach. *ISME J* 2012;
674 **6**: 927-38.
675 28. Plichta DR, Juncker AS, Bertalan M, Rettedal E, Gautier L, Varela E, et al. Transcriptional
676 interactions suggest niche segregation among microorganisms in the human gut. *Nat Microbiol* 2016; **1**:
677 16152.
678 29. Root RB. The niche exploitation pattern of the blue-gray gnatcatcher. *Ecol Monogr* 1967; **37**:
679 317-50.
680 30. Burke C, Steinberg P, Rusch D, Kjelleberg S, Thomas T. Bacterial community assembly based on
681 functional genes rather than species. *Proc Natl Acad Sci USA* 2011; **108**: 14288-93.
682 31. McGill B, Enquist B, Weiher E, Westoby M. Rebuilding community ecology from functional
683 traits. *Trends Ecol Evol* 2006; **21**: 178-85.
684 32. Bhatnagar JM, Peay KG, Treseder KK. Litter chemistry influences decomposition through
685 activity of specific microbial functional guilds. *Ecol Monogr* 2018; **88**: 429-44.
686 33. Prosser JI. Dispersing misconceptions and identifying opportunities for the use of 'omics' in soil
687 microbial ecology. *Nat Rev Microbiol* 2015; **13**: 439-46.
688 34. Jaeger CH, Lindow SE, Miller W, Clark E, Firestone MK. Mapping of sugar and amino acid
689 availability in soil around roots with bacterial sensors of sucrose and tryptophan. *Appl Environ Microbiol*
690 1999; **65**: 2685-90.
691 35. Placella SA, Brodie EL, Firestone MK. Rainfall-induced carbon dioxide pulses result from
692 sequential resuscitation of phylogenetically clustered microbial groups. *Proc Natl Acad Sci USA* 2012;
693 **109**: 10931-6.
694 36. Fierer N, Jackson RB. The diversity and biogeography of soil bacterial communities. *Proc Natl*
695 *Acad Sci USA* 2006; **103**: 626-31.

37. Griffiths RI, Whiteley AS, O'Donnell AG, Bailey M. Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Appl Environ Microbiol* 2000; **66**: 5488-91.

38. Barnard RL, Osborne CA, Firestone MK. Responses of soil bacterial and fungal communities to extreme desiccation and rewetting. *ISME J* 2013; **7**: 2229-41.

39. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012; **9**: 671-5.

40. Rinke C, Lee J, Nath N, Goudeau D, Thompson B, Poulton N, et al. Obtaining genomes from uncultivated environmental microorganisms using FACS–based single-cell genomics. *Nature Protocols* 2014; **9**: 1038-48.

41. Tremblay J, Singh K, Fern A, Kirton ES, He S, Woyke T, et al. Primer and platform effects on 16S rRNA tag sequencing. *Front Microbiol* 2015; **6**: 771.

42. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* 2011; **108**: 4516-22.

43. White TJ, Bruns T, Lee S, Taylor J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis M, Gelfland D, Sninsky J, White T, editors. PCR protocols: A guide to methods and applications. San Diego, CA: Academic Press; 1990. p. 315-22.

44. Ihrmark K, Bödeker ITM, Cruz-Martínez K, Friberg H, Kubartova A, Schenck J, et al. New primers to amplify the fungal ITS2 region – evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiol Ecol* 2012; **82**: 666-77.

45. Li M, Copeland A, Han J. DUK - A fast and efficient kmer based sequence matching tool. *LBNL-4516E-Poster* 2011.

46. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012; **9**: 357-9.

721 47. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA
722 gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2012; **41**: D590-
723 D6.
724 48. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, et al. An improved
725 Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and
726 archaea. *ISME J* 2012; **6**: 610-8.
727 49. Chan PP, Lowe TM. GtRNAdb: a database of transfer RNA genes detected in genomic sequence.
728 *Nucleic Acids Res* 2009; **37**: D93-7.
729 50. Abe T, Ikemura T, Sugahara J, Kanai A, Ohara Y, Uehara H, et al. tRNADB-CE 2011: tRNA
730 gene database curated manually by experts. *Nucleic Acids Res* 2011; **39**: D210-3.
731 51. Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome
732 assemblies. *Bioinformatics* 2011; **27**: 2957-63.
733 52. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;
734 **26**: 2460-1.
735 53. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of
736 rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 2007; **73**: 5261-7.
737 54. Bushnell B. BBTools Software Package; sourceforge.net/projects/bbmap/
738 55. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A new
739 genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455-
740 77.
741 56. Tennessen K, Andersen E, Clingenpeel S, Rinke C, Lundberg DS, Han J, et al. ProDeGe: a
742 computational protocol for fully automated decontamination of genomes. *ISME J* 2016; **10**: 269-72.
743 57. Starr EP, Shi S, Blazewicz SJ, Probst AJ, Herman DJ, Firestone MK, et al. Stable isotope
744 informed genome-resolved metagenomics reveals that *Saccharibacteria* utilize microbially-processed
745 plant-derived carbon. *Microbiome* 2018; **6**: 122.

746 58. Zhalnina K, Louie KB, Hao Z, Mansoori N, da Rocha UN, Shi S, et al. Dynamic root exudate
747 chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community
748 assembly. *Nat Microbiol* 2018; **3**: 470-80.
749 59. Probst AJ, Ladd B, Jarett JK, Geller-McGrath DE, Sieber CMK, Emerson JB, et al. Differential
750 depth distribution of microbial function and putative symbionts through sediment-hosted aquifers in the
751 deep terrestrial subsurface. *Nat Microbiol* 2018; **3**: 328-36.
752 60. Raes J, Korbel JO, Lercher MJ, von Mering C, Bork P. Prediction of effective genome size in
753 metagenomic samples. *Genome biology* 2007; **8**: R10.
754 61. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene
755 recognition and translation initiation site identification. *BMC Bioinformatics* 2010; **11**: 119.
756 62. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated
757 carbohydrate-active enzyme annotation. *Nucleic Acids Res* 2012; **40**: W445-W51.
758 63. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*
759 2000; **28**: 27-30.
760 64. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides
761 from transmembrane regions. *Nat Methods* 2011; **8**: 785-6.
762 65. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq
763 data with DESeq2. *Genome biology* 2014; **15**: 550.
764 66. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME
765 allows analysis of high-throughput community sequencing data. *Nat Methods* 2010; **7**: 335-6.
766 67. Team RC. R: A language and environment for statistical computing: R Foundation for Statistical
767 Computing 2017; http://www.R-project.org/
768 68. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. vegan: Community
769 Ecology Package. R package version 2.0-6 2013; http://CRAN.R-project.org/package=vegan
770 69. Berlemont R, Martiny AC. Genomic Potential for Polysaccharide Deconstruction in Bacteria.
771 *Appl Environ Microbiol* 2015; **81**: 1513-9.

772 70. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, et al. A
773 standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat*
774 *Biotechnol* 2018; **36**: 996.
775 71. Rocha DJP, Santos CS, Pacheco LGC. Bacterial reference genes for gene expression studies by
776 RT-qPCR: survey and analysis. *Antonie Van Leeuwenhoek* 2015; **108**: 685-93.
777 72. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a
778 better web interface. *Nucleic Acids Res* 2008; **36**: W5-W9.
779 73. Moran MA, Satinsky B, Gifford SM, Luo H, Rivers A, Chan L-K, et al. Sizing up
780 metatranscriptomics. *ISME J* 2013; **7**: 237-43.
781 74. Yergeau E, Sanschagrin S, Maynard C, St-Arnaud M, Greer CW. Microbial expression profiles in
782 the rhizosphere of willows depend on soil contamination. *ISME J* 2014; **8**: 344-58.
783 75. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-
784 throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 2012; **6**:
785 1621-4.
786 76. Schimel JP, Schaeffer SM. Microbial control over carbon cycling in soil. *Front Microbiol* 2012;
787 **3**: 1-11.
788 77. Martiny JBH, Jones SE, Lennon JT, Martiny AC. Microbiomes in light of traits: A phylogenetic
789 perspective. *Science* 2015; **350**: aac9323-1-8.
790 78. Moorhead DL, Sinsabaugh RL. A theoretical model of litter decay and microbial interaction. *Ecol*
791 *Monogr* 2006; **76**: 151-74.
792 79. Goldford JE, Lu N, Bajić D, Estrela S, Tikhonov M, Sanchez-Gorostiaga A, et al. Emergent
793 simplicity in microbial community assembly. *Science* 2018; **361**: 469-74.
794 80. Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, et al. Thousands of
795 microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat*
796 *Commun* 2016; **7**: 13219.

797 81. Diamond S, Andeer P, Li Z, Crits-Christoph A, Burstein D, Anantharaman K, et al. Processing of
798 grassland soil C-N compounds into soluble and volatile molecules is depth stratified and mediated by
799 genomically novel bacteria and archaea. *bioRxiv* 2018: 445817.
800 82. Shi SJ, Nuccio EE, Shi ZJ, He ZL, Zhou JZ, Firestone MK. The interconnected rhizosphere: High
801 network complexity dominates rhizosphere assemblages. *Ecol Lett* 2016; **19**: 926-36.
802 83. Macarthur R, Levins R. The limiting similarity, convergence, and divergence of coexisting
803 species. *Am Nat* 1967; **101**: 377-85.
804 84. Kuzyakov Y, Friedel J, Stahr K. Review of mechanisms and quantification of priming effects.
805 *Soil Biol Biochem* 2000; **32**: 1485-98.
806 85. Fontaine S, Mariotti A, Abbadie L. The priming effect of organic matter: a question of microbial
807 competition? *Soil Biol Biochem* 2003; **35**: 837-43.
808 86. Pascault N, Ranjard L, Kaisermann A, Bachar D, Christen R, Terrat S, et al. Stimulation of
809 different functional groups of bacteria by various plant residues as a driver of soil priming effect.
810 *Ecosystems* 2013; **16**: 810-22.
811 87. van Passel MWJ, Kant R, Palva A, Copeland A, Lucas S, Lapidus A, et al. Genome sequence of
812 the Verrucomicrobium *Opitutus terrae* PB90-1, an abundant inhabitant of rice paddy soil ecosystems. *J*
813 *Bacteriol* 2011; **193**: 2367-8.
814 88. Wertz JT, Kim E, Breznak JA, Schmidt TM, Rodrigues JLM. Genomic and physiological
815 characterization of the Verrucomicrobia isolate *Diplosphaera colitermitum* gen. nov., sp. nov., reveals
816 microaerophily and nitrogen fixation genes. *Appl Environ Microbiol* 2012; **78**: 1544-55.
817 89. Lin JY, Russell JA, Sanders JG, Wertz JT. *Cephaloticoccus* gen. nov., a new genus of
818 'Verrucomicrobia' containing two novel species isolated from *Cephalotes* ant guts. *Int J Syst Evol*
819 *Microbiol* 2016; **66**: 3034-40.
820 90. Hünninghaus M, Dibbern D, Kramer S, Koller R, Pausch J, Schloter-Hai B, et al. Disentangling
821 carbon flow across microbial kingdoms in the rhizosphere of maize. *Soil Biol Biochem* 2019; **134**: 122-
822 30.

823 91. Montgomery L, Flesher B, Stahl D. Transfer of *Bacteroides succinogenes* (Hungate) to
824 *Fibrobacter* gen. nov. as *Fibrobacter succinogenes* comb. nov. and description of *Fibrobacter intestinalis*
825 sp. nov. *Int J Syst Bacteriol* 1988; **38**: 430-5.
826 92. Ivanova AA, Wegner CE, Kim Y, Liesack W, Dedysh SN. Identification of microbial populations
827 driving biopolymer degradation in acidic peatlands by metatranscriptomic analysis. *Mol Ecol* 2016; **25**:
828 4818-35.
829 93. Rahman NA, Parks DH, Vanwonterghem I, Morrison M, Tyson GW, Hugenholtz P. A
830 phylogenomic analysis of the bacterial phylum Fibrobacteres. *Front Microbiol* 2016; **6**: 1469.
831 94. Edwards J, Johnson C, Santos-Medellín C, Lurie E, Podishetty NK, Bhatnagar S, et al. Structure,
832 variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci USA* 2015; **112**:
833 E911-E20.
834 95. Pitcher RS, Brittain T, Watmugh NJ. Cytochrome cbb3 oxidase and bacterial microaerobic
835 metabolism. *Biochem Soc Trans* 2002; **30**: 653-8.
836 96. Højberg O, Sørensen J. Microgradients of microbial oxygen consumption in a barley rhizosphere
837 model system. *Appl Environ Microbiol* 1993; **59**: 431-7.
838 97. Keiluweit M, Bougoure JJ, Nico PS, Pett-Ridge J, Weber PK, Kleber M. Mineral protection of
839 soil carbon counteracted by root exudates. *Nat Clim Change* 2015; **5**: 588-95.
840 98. López-Mondéjar R, Brabcová V, Štursová M, Davidová A, Jansa J, Cajthaml T, et al.
841 Decomposer food web in a deciduous forest shows high share of generalist microorganisms and
842 importance of microbial biomass recycling. *ISME J* 2018; **12**: 1768-78.
843 99. Blazewicz SJ, Schwartz E, Firestone MK. Growth and death of bacteria and fungi underlie
844 rainfall-induced carbon dioxide pulses from seasonally dried soil. *Ecology* 2014; **95**: 1162-72.
845 100. Li J, Mau R, Dijkstra P, Koch B, Schwartz E, Purcell A, et al. Genomic traits predict microbial
846 growth in culture but fail in soils, except during resource pulses. *ISME J*; **(in press)**.
847 101. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, et al. A new view of the
848 tree of life. *Nat Microbiol* 2016; **1**: 16048.

849