Transcription of protease and chitinase genes provides a window onto macromolecular organic nitrogen decomposition in soil

Ella T. Sieradzki[1], Erin E. Nuccio[2], Jennifer Pett-Ridge[2], Mary K. Firestone[1,3#]

[1]Department of Environmental Science, Policy and Management, University of California, Berkeley, California, USA
[2]Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, California, USA
[3] Climate & Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

[#]Address correspondence to: mkfstone@berkeley.edu

Abstract

Nitrogen is a common limiting nutrient in soil in part because most N is present as macromolecular organic compounds, not directly available to plants. The microbial community present in soil near roots (rhizosphere) is in many ways analogous to the human gut microbiome, transforming nutrients present in organic substrates to forms available to plants through extracellular enzymes. Many recent studies have focused on the genetic potential for nitrogen cycling by bacteria in the rhizosphere, and on measuring inorganic N pools and fluxes. Between those two bodies of knowledge, there is scarce information on functionality of macromolecular nitrogen decomposing bacteria and fungi and how it relates to life stages of the plant. This is particularly important as many soil bacteria identified in community composition studies can be inactive or not viable. Here we use a time-series of metatranscriptomes from rhizosphere and bulk soil bacteria and fungi to follow extracellular protease and chitinase expression during rhizosphere aging. In addition, we explore the effect of adding plant litter as a source of macromolecular carbon and nitrogen. Expression of extracellular proteases increased over time in the absence of litter, more so in the presence of roots, whereas the dominant chitinase (*chit1*) was upregulated with exposure to litter. Structural groups of proteases were surprisingly dominated by serineproteases, possibly due to the importance of betaproteobacteria and actinobacteria in this grassland soil. Extracellular proteases of betaprotebacterial origin were more highly expressed in the presence of roots, whereas deltaroteobacteria and fungi responded to the presence of litter. We found functional guilds specializing in decomposition of proteins in the rhizosphere, detritusphere and in the vicinity of aging roots. We also identify a guild that appears to specialize in protein decomposition in the presence of roots and litter and increases its activity in aging rhizosphere, which may imply that this guild targets rhizodeposits or the senescing root itself as a protein source. Different temporal patterns of guilds imply that rather than functional redundancy, microbial decomposers operate within distinct niches.

Keywords

Rhizosphere, microbiome, extracellular protease, functional guilds

Background

Plants are commonly limited by nitrogen in temperate soils as they cannot directly access macromolecular nitrogen, the most abundant pool of nitrogen in soil [1–7]. Therefore, they depend on microbial degradation of macromolecular forms of nitrogen such as proteins and chitin [8–10] and potentially influence this degradation by priming [11–14]. The current paradigm is that depolymerization of high molecular weight organic nitrogen is the rate limiting step in nitrogen mineralization in soil [15, 16]. While rates of nitrogen mineralization and the genetic potential for macromolecular nitrogen degradation have been studied, data linking the two via gene expression is currently scarce.

The rhizosphere is a hotspot for microbial biomass and activity due to root exudation and debris, and the microbes present are critical to plant growth [17]. However, root exudates, which serve as a carbon source for rhizosphere bacteria, decline near older root sections [18]. Therefore, it has been hypothesized that bacteria in a mature rhizosphere environment turn to less labile, higher C/N sources of carbon such as plant litter. An increased demand for N is

consistent with the higher rates of gross nitrogen mineralization and ammonium consumption found in the rhizosphere of mature compared to young *Avena sp*. [19].

We have previously found that soil bacteria, in the rhizosphere and in the presence of plant litter, act within functional guilds defined by expression of carbohydrate active enzymes (CAZy) [20]. We postulated that a analogous functional organization might occur in expression of extracellular proteases, as these diverse enzymes have also been shown to target a variety of substrates and be influenced by soil characteristics [21–23]. It has been shown that activity of extracellular proteases drives soil N cycling (Schimel and Bennet 2004), but what controls their activity is not always clear (Brzostek and Finzi 2011). To begin elucidating these mechanisms, we investigated community-wide expression and taxon-specific differential expression of extracellular proteases and chitinases in the rhizosphere vs. bulk soil over time and in the presence of root litter.

We analyzed metatranscriptomes from a robust, well developed experimental approach [19, 20, 24, 25] based on a highly characterized annual grassland system in Northern California. The wild oat *Avena fatua* [2] was grown in "native" soil to elucidate the role of diverse indigenous rhizosphere bacteria in macromolecular nitrogen degradation by extracellular proteases and chitinases throughout root growth and aging. We find that expression of extracellular protease genes is taxonomically and structurally specific in the presence of roots and/or root detritus. In addition, we compare and contrast functional guilds based on extracellular protease expression with guilds defined by expression of carbohydrate active enzymes (CAZy) capable of degrading root litter.

Methods

Experimental design, sample collection and sequencing

The experimental design, sample collection and sequence data processing were described in a recent publication [20]. Briefly, *Avena fatua* plants were grown in microcosms containing Bearwallow–Hellman loam (pH 5.6, 2% total C) with a sidecar for six weeks, after which the sidecar was made available for roots to grow into it. As roots grew into the side car, their growth timeline was marked on the transparent wall of the sidecar at 3 days, 6 days, 12 days and 22 days. Half of the sidecars contained dried *A. fatua* root detritus amendment (C:N = 13.4). Bulk soil bags, inaccessible to roots, were placed in each sidecar and were either amended with detritus or not. At each timepoint paired rhizosphere and bulk soil triplicates were destructively harvested from amended and unamended microcosms for a total of 48 samples representing four treatments: rhizosphere, rhizosphere + litter, bulk, and bulk + litter.

Harvested soil (1 g) was placed immediately in Lifeguard Soil Preservation Reagent (MoBio) and processed according to the company protocol. Roots and supernatant were removed, and the soil was stored in -80°C.

DNA/RNA were co-extracted using phenol-chloroform and separated with an AllPrep kit (Qiagen). RNA was DNase treated (TURBO DNase, Thermo-Fisher Scientific), depleted from ribosomal RNA (RiboZero, Illumina) and reverse transcribed. Metatranscriptomes were sequenced from each sample on Illumina HiSeq 2000 2x150 (TruSeq SBS v3).

Expression of genes identified in assembled metatranscriptomes

Raw reads were quality-trimmed (Q20) and rRNA and tDNA were removed. Library size, evenness, richness and Shannon diversity were comparable between experimental groups with a mean library size of 43 M paired end reads. Quality-controlled metatranscriptomic reads were assembled into contigs within each sample [20]. Contigs smaller than 200bp were discarded and the remaining contigs from all samples were clustered at 99% id with cd-hit-est [26]. ORFs were predicted using Prodigal [27]. Extracellular protease ORFs were identified by reciprocal blast to extracellular peptidases from the MEROPS database [28, 29]. ORFs were assigned a peptidase group (serine-, metallo-, cysteine-peptidase and others) by their best reciprocal blast hit. Taxonomy was determined by best blast hit to the NCBI non-redundant database (NR, access date July 29th 2019). Reads were then mapped back to ORFs requiring minimum identity 95% and 75% breadth using bbmap [30]. Read counts were normalized to sequencing depth using DESeq2. Heat maps of normalized counts were generated in R using the heatmap2 function in gplots [31]. Normalized transcript counts per gene per time point were compared between groups (rhizosphere, litter and litter-amended rhizosphere) using the Kruskal-Wallis test. Boxplots were generated in ggplot2 [32].

Expression of extracellular protease genes from a curated collection of Hopland-soil genomes

The quality-controlled reads were mapped using BBsplit [30] requiring 80% identity against a dereplicated reference database of 282 Hopland soil genomes including isolates [14], single amplified genomes (SAGs) [20], metagenomic assembled genomes (MAGs, NCBI PRJNA517182) and stable isotope probing MAGs (SIP-MAGs) [33]. On average, 12.3% (range 6.2-31.9%) of the reads per library mapped unambiguously to the reference database.

Open reading frames (ORFs) in the genomes were predicted in all genomes using prodigal [27] and annotated using KEGG [34] and ggKbase (http://ggkbase.berkeley.edu). Extracellular proteases, as they do not have hidden Markov models (HMMs) capable of separating them from intracellular proteases, were identified by gene nucleotide identity of at least 90% and coverage of at least 60% to extracellular protease ORFs from the metatranscriptomes (see below). Gene counts were identified using Rsubread featureCounts [35].

All features and their abundance (represented by metatranscriptomic read counts normalized to sequencing depth) were analyzed with DESeq2 [36] requiring p-value < 0.05 (adjusted for multiple comparisons). Ordination and visualization were conducted in R using ggplot2 [32] and vegan [37], and the vegan function Adonis was used to detect significant factors affecting expression of nitrogen cycling genes. Guilds were assigned by hierarchical clustering based on differential expression of extracellular proteases compared to unamended bulk soil, generating 4 functional guilds: “Late synergists”, “Rhizosphere”, “Detritusphere”, and “Aging root”.

Phylogenomic analysis of the MAGs was performed using GTDB-tk [38]. Verification of the taxonomic assignment of MAG Burkholderiales_62_29 was done using GToTree [39] with complete reference genomes of Betaproteobacteria from RefSeq (Feb 28, 2020).

Results

Expression of enzymes targeting macromolecular N

Transcripts for 4924 protease genes from bacteria and fungi were quantified in the sampled soils. Overall, extracellular protease expression was significantly affected by time (p=0.038), litter amendment (p=1.1e-27) as well as interactions between time:treatment (2.3e-17), time:location (p=3.9e-5) and location:treatment (p=8.2e-5) (ANOVA, p adjusted for multiple comparisons). Expression in unamended soil (no litter) generally increased over time (fig. 1A), whereas amended soil displayed a more complex pattern (fig. 1B). The differences in these responses appears to be due to taxonomic clades being affected differently by the presence of roots and litter. In contrast, the 73 distinct transcripts detected from chitinase genes were expressed at a much lower level compared to extracellular proteases and had consistent significant effects throughout the experiment. Interestingly, the expression of the *chit1* gene was higher in litter-amended samples (p=2e-11) (fig. 1C). Transcripts for chitinase KEGG orthologs *CHI3L1_2*, *CHI3L1_4* and *chiA* were not detected at all and putative chitinase K03791 had low expression and no significant effects.

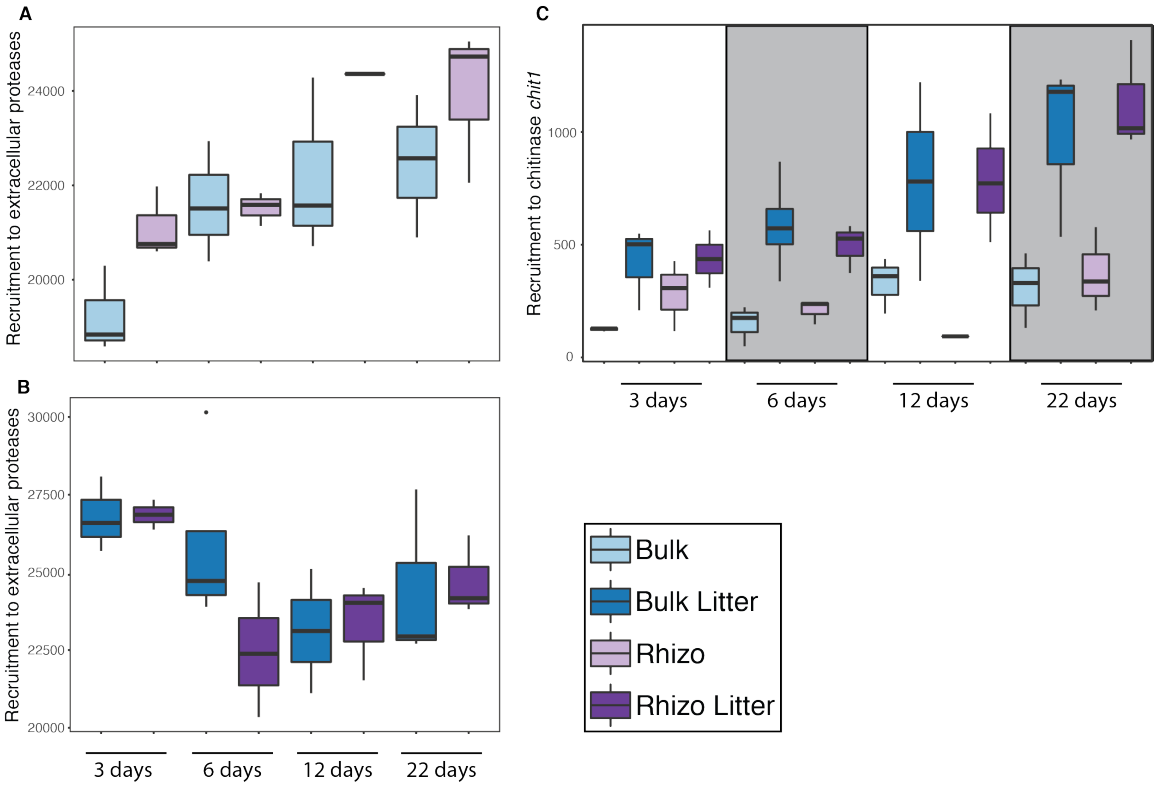

Figure 1: Aggregated expression of extracellular proteases without litter amendment (A), extracellular proteases with litter amendment (B) and chitinase gene *chit1* (C). The legend applies to all panels. Boxplots represent 75% of the data, whiskers denote 90% of the data and dots (e.g. panel A bulk litter 6 days) represent outliers. Note that the scale of the Y axis varies between panels.

Structural groups of extracellular proteases

The main groups of extracellular proteases found in the assembled metatranscriptomes were serine-, metallo- and cysteine-proteases in decreasing order (2679, 1949 and 496 out of

5295 variants clustered at 99% identity respectively). Expression patterns of all variants in each group reflected the same order. Expression of serine- was higher than metallo-protease within treatment. Expression of both serine- and metallo-proteases was significantly greater in the presence of litter compared to no litter (Kruskal-Wallis $p<2.2e-16$, Dunn test with BH correction $p<0.05$) (fig. 2).

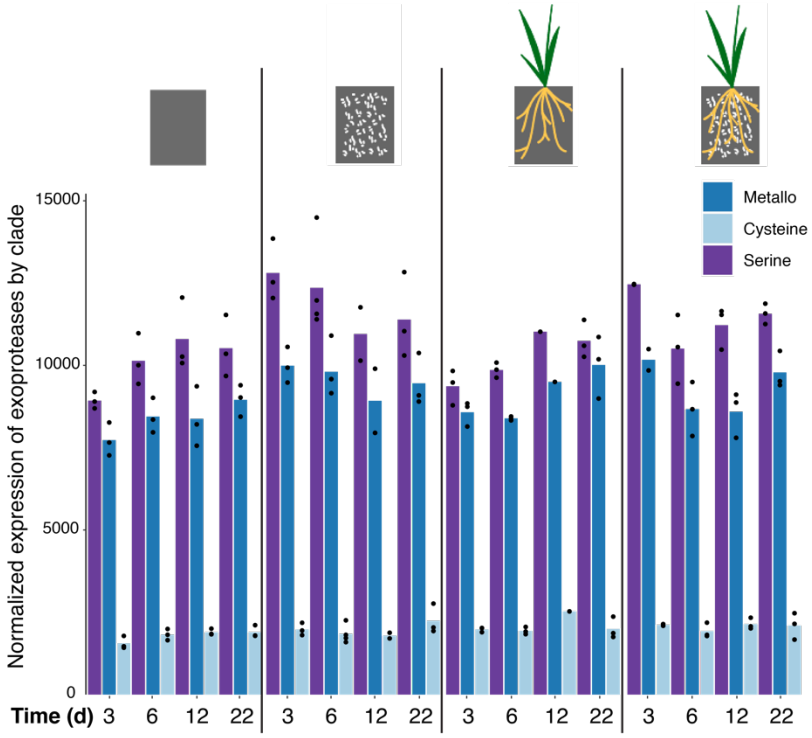

Figure 2: Aggregated normalized expression of extracellular protease structural groups over time. Normalized expression per replicate is plotted as dots on top of the bars. Experimental groups shown per time point are (left to right): unamended bulk soil, litter-amended bulk soil, unamended rhizosphere and litter-amended rhizosphere. Extracellular protease groups are (left to right): serine-, metallo- and cysteine-protease.

Taxonomy of extracellular proteases

ORFs of extracellular proteases were assigned taxonomy by blastp best hit against the Non-redundant database (NR). As extracellular proteases are not marker genes, we considered taxonomic assignments only at the order level or higher, with the exception of *Rhizobacter*, for which amino acid percent identity values were extremely high (see below).
Hierarchical clustering based on differential expression of the most highly upregulated ORFs compared to unamended bulk soil revealed clear rhizosphere and detritusphere guilds (sup. fig. S1).

In the rhizosphere, we found that Betaproteobacterial exoproteases were significantly upregulated (fig. 3A). Of 547 variants of exoproteases from this class, 442 were assigned as Burkholderiales. In contrast, Cyanobacteria were significantly downregulated in the rhizosphere

compared to bulk soil, and Thaumarchaeota trended similarly (fig. 3B), though this trend was not statistically significant likely due to time point 3 (12 days).

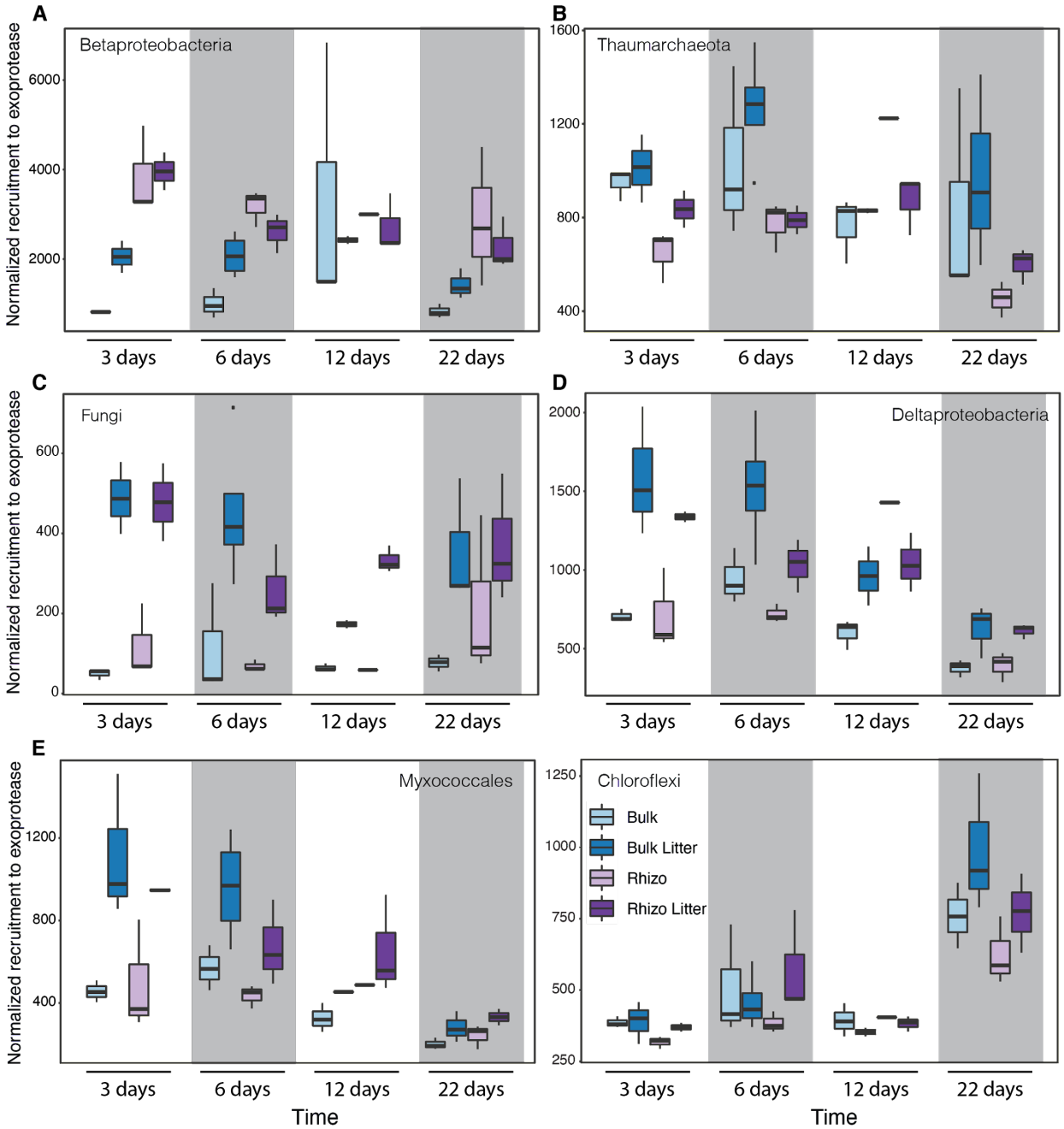

Figure 3: Normalized expression of extracellular proteases over time within select taxonomic groups: (A) Betaprotebacteria (n=547) were upregulated in the rhizosphere, (B) Thaumarchaeota (n=112) were downregulated in the rhizosphere, (C) Fungi (n=99) were upregulated in the presence of litter, (D) Detaproteobacteria (n=228) were upregulated in the presence of litter, (E) Myxococcales (n=147) were upregulated in the presence of litter and decreased over time and (F) Chloroflexi (n=153) increased in the last time point. The legend in panel F applies to all panels.

Several taxonomic groups were significantly upregulated (adjusted p-value < 0.1) in the presence of litter: kingdom Fungi (fig. 3C), class Deltaproteobacteria (fig. 3D), as well as highly represented phyla Bacterioidetes and Verrucomicrobia (sup. table S1) and classes Chitinophagia and Gammaproteobacteria (sup. table S2).

An interesting temporal trend was revealed in predatory bacteria Myxococcales (fig. 3E) as well as Bdellovibrionales and Cytophagia, all of which decreased expression over time. Inversely, Some clades such as Chloroflexi (fig. 3F) and Actinobacteria increased expression of exoproteases in the last time point.

Actinobacteria and Acidobacteria, both of which had a high number of variants and high expression of exoproteases, did not display a preference for location or amendment. Normalized expression by phylum, class and order are summarized in sup. fig. S2, sup. fig. S3, sup. fig. S4 and sup. table S1, sup. table S2 and sup. table S3 respectively.

Functional guilds

In order to define functional guilds, we mapped transcriptomic reads to a collection of genomes from soil from the same site (Nuccio et al. 2020). Read counts were used to determine differential expression compared to unamended bulk soil using DESeq2. Each genome contained multiple genes coding for extracellular proteases. Hierarchical clustering of the mean differential expression of extracellular protease genes within each genome was used to define four functional guilds: “rhizosphere”, “detritusphere”, “aging root” and “late synergists” (fig. 4). Differential expression of extracellular proteases in late synergists increased at the last time point in the presence of both roots and litter more than would have been predicted by the sum of the treatments alone. This guild is comprised of a single MAG named Burkholderiales_62_29. Two independent phylogenomic analyses identified this MAG as *Rhizobacter* (sup. fig. S5). Burkholderiales_62_29 has 16 different extracellular protease genes with similar significant upregulation patterns.

Three of the guilds, namely rhizosphere, detritusphere and aging root, were previously identified by a hierarchical clustering based on differential expression of carbohydrate-active enzymes (CAZy) from the same dataset [20]. Several of the taxa identified, Burkholderiales 54_8, *Massilia* 66_12 and *Janthinobacterium* 65_15 were assigned to the rhizosphere guild by both CAZy and extracellular protease expression. *Devosia* JGI 030 and *Asticcacaulis* 63_11 were members in both detritusphere guilds. Finally, both aging root guilds contained several *Streptomyces* genomes: 70_12, 71_16, 71_13 and 70_9. Note that the taxa comprising the aging root guild also have extracellular protease upregulation in the detritusphere. “Late synergists” member Burkholderiales bin 62_29 was assigned to the rhizosphere guild by CAZy expression. However, by extracellular protease expression it clustered closer to the aging root guild (fig. 4).

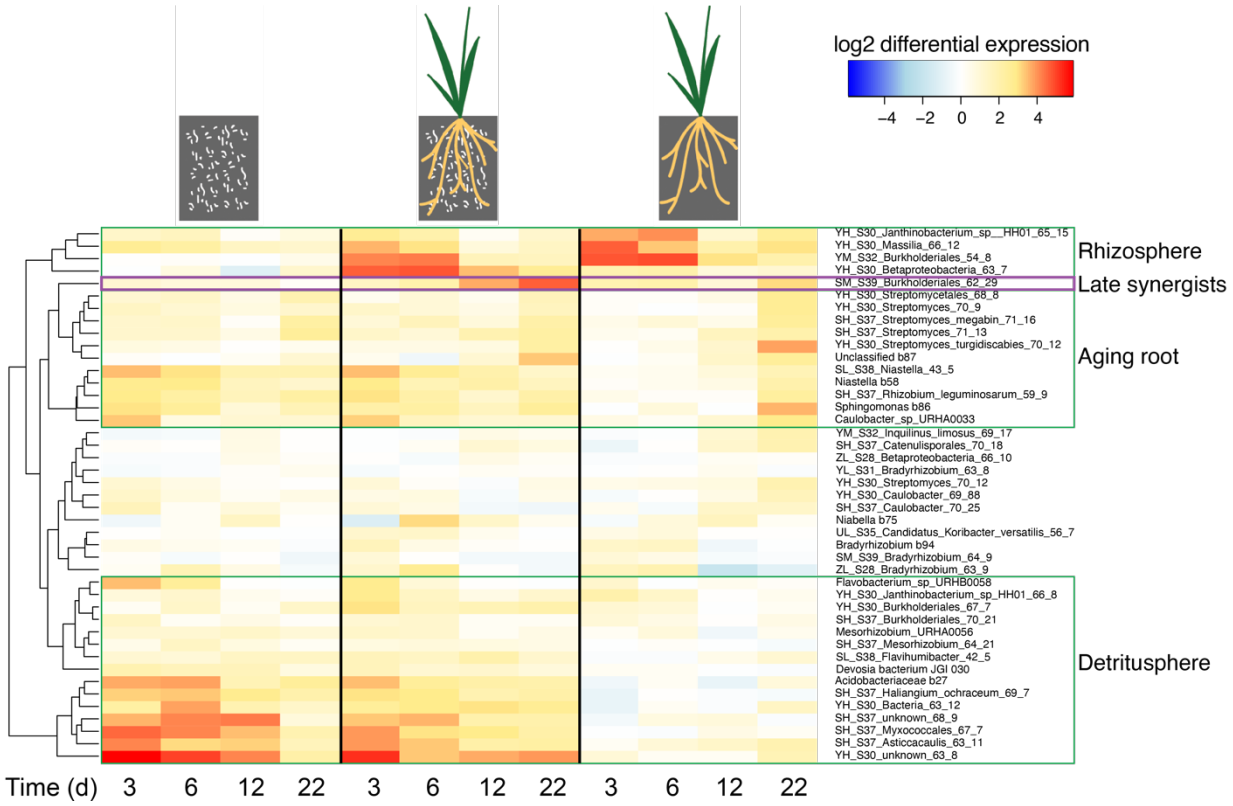

Figure 4: Functional guilds defined by hierarchical clustering of extracellular protease differential expression. Each row represents the mean differential expression of extracellular protease genes within a genome. Groups are (left to right): litter-amended bulk soil, litter-amended rhizosphere and unamended rhizosphere. Time points are indicated on the bottom in days. Note that a genome may contain more than one exoprotease gene and that reads were mapped loosely, therefore each genome also represents closely related taxa. Differential expression values per gene that were not statistically significant were converted to zero (0) before averaging.

Nitrogen cycling genes in Rhizobacter

As the late synergists guild was comprised of taxa closely related to *Rhizobacter*, we explored extracellular protease ORFs assigned as *Rhizobacter* in the metatranscriptome assemblies, as well as all nitrogen cycling pathways identified in the *Rhizobacter* genome (Burkholderiales_62_29). The best hit for most assembled ORFs (84±8.4% amino acid identity) was to *Rhizobacter* from a root metagenome [40] (fig. 5A). Expression of N cycling genes identified in the *Rhizobacter* genome, i.e. ammonium transporters, ammonia assimilation (GS/GOGAT; *glnA, gltBD*) and most of its regulatory pathway activated under N limitation (*glnDLG*), were significantly upregulated in amended rhizosphere after 22 days (fig. 5B). Most of these genes were also significantly upregulated in unamended rhizosphere after 22 days, but to a lesser extent.

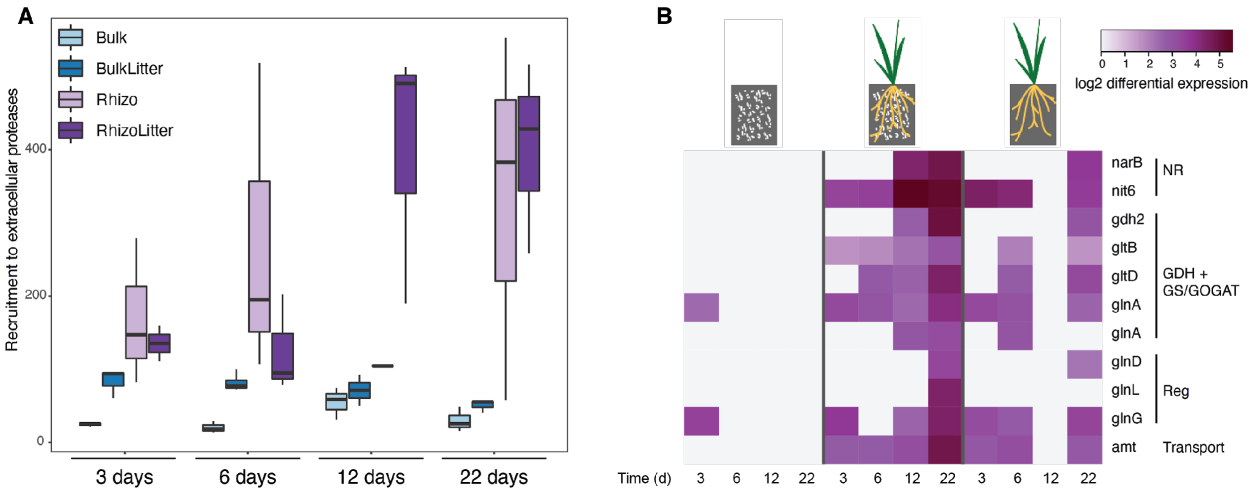

Figure 5: Expression of *Rhizobacter* nitrogen cycling genes. (A) Extracellular proteases from the assembled metatranscriptomes over time. (B) Log-transformed differential expression of nitrogen cycling genes identified in the *Rhizobacter* MAG (Burkholderiales_62_29) compared to unamended bulk soil. Pathways are indicated on the right: NR = assimilatory nitrate reduction, GDH = glutamate dehydrogenase, GS/GOGAT = glutamate synthase, Reg = regulation of GS/GOGAT, Transport = ammonium transporters. This genome has two copies of *glnA*.

Discussion

Nitrogen in grassland soil is present mostly in the form of organic macromolecules such as proteins and chitin [2, 41]. There is evidence that rhizosphere bacteria can acquire nitrogen through degradation of extracellular macromolecular N, and that this process can be primed by root exudates especially in unfertilized or N limited soil [7, 13, 42]. Amino acids resulting from protein degradation are taken up by microorganisms, and roughly 30% of the amino-acid carbon is respired off, leading to excretion of ammonium which could benefit the plant (Hill and Jones 2019; Roberts et al. 2009). Chitinase expression was two orders of magnitude lower than that of proteases, implying a lesser role for chitin as an N source in our system, assuming no significant difference in downstream processing. Alternatively, it could simply mean that chitinases are more stable enzymes that do not require the same frequency of transcription and translation. The increase over time in expression of chitinase in litter amended samples could imply that the litter attracted arthropods and saprotrophic fungi, which in turn were decomposed.

Community-wide expression of all extracellular proteases revealed an increase over time in unamended rhizosphere soil. In litter amended rhizosphere, peak transcription was in three days old roots (first sampling time), but the lower levels of expression at 6 days were followed by an increasing trend from the 6 days on. This suggests that either this litter could serve as a nitrogen source for rhizosphere bacteria, or that the carbon acquired by degrading the litter created a higher demand for nitrogen which could not be fulfilled by root exudates. In a previous study of *Avena fatua*, rhizosphere protease specific activity (by soil enzyme activity assay) was significantly higher in rhizosphere compared to bulk soil (no litter amendments included). This protease activity accounted for a third of the mineralized N [24]. The results from our unamended soils agree with these findings. Moreover, the increase in expression over time in the rhizosphere lends support to a long-standing hypothesis that as the rhizosphere ages it becomes

more nitrogen limited; thus organic nitrogen would need to be mineralized to support the microbial community.

Like any chemical reaction, macromolecular N depolymerization can be regulated by substrate availability and end-product concentration. At first, litter amendment led to high expression of extracellular proteases, suggesting that this expression is substrate dependent [43, 44] and linking carbon and nitrogen cycling [45]. However, protease activity in grassland soil has also been shown to be affected by a multitude of parameters, such as end-product concentration, above ground net primary productivity [23] and soil aggregation [22]. In the later time points the expression of extracellular proteases in litter amended samples decreased. This could be attributed to a decrease in the easily accessible substrate due to litter degradation earlier on. Surprisingly, there was also an interesting increase over time in expression in unamended samples although no nitrogen or carbon sources were added. We hypothesize that this increase over time may be related to depletion of end products, leading soil bacteria to invest more in expressing depolymerization enzymes in order to access macromolecular nitrogen. We also observed higher variability in bulk soil expression of extracellular proteases. This could potentially be due to the more heterogeneous nature of bulk soil compared to the rhizosphere, which is spatially limited and highly affected by roots and therefore possibly more uniform in terms of C-availability.

Spatiotemporal patterns of differential expression of assembled ORFs and of specific genomes were used to identify functional guilds: rhizosphere, detritusphere, aging root and late synergists. These microbial guilds underwent distinct successional patterns that appear to be driven by niche conditions. Nuccio and coworkers (2020) hypothesized that at 22 days, the senescing roots began to elicit the same transcription response as root detritus. Composition of the microbial community in each experimental group did not change detectably over the course of this experiment [20]. Therefore, the changing expression patterns over time must be related to changes in environmental conditions or cues, such as macromolecular N availability and inorganic N availability. Moreover, while genomes containing similar enzymes may create the impression of functional redundancy, these temporal patterns indicate that this is not the case. Rather, different taxa perform a similar function under different conditions. Finally, we found temporal patterns of expression in predatory bacteria which could indirectly contribute to nitrogen mineralization in young rhizosphere through predation.

There is some overlap between guild members we identified here, and similar functional guilds previously defined by expression of carbohydrate active enzyme (CAZy) [20]. The rhizosphere guild was comprised of betaproteobacteria known to be associated with rhizosphere soil [25, 46, 47] such as Burkholderiales, *Janthinobacterium* and *Massilia*. In contrast, the detritusphere guild was more taxonomically diverse. The upregulation of carbon depolymerization may require accessing nitrogen through expression of extracellular proteases under the same conditions, leading to an overlap in guild membership. Such a causative correlation has been previously implied in loam agricultural soil [45].

We identified a "specialized guild" comprised of *Rhizobacter spp.* which increased upregulation of extracellular proteases over time in rhizosphere amended with litter. In addition to upregulating extracellular proteases, this *Rhizobacter* also upregulated various mechanisms for uptake of inorganic nitrogen which could result from mineralization (e.g. assimilatory nitrate reduction, ammonium assimilation, ammonium transporters and activators of GS/GOGAT under N limitation). This *Rhizobacter* MAG has been proposed to represent a plant pathogen due to the presence of type 3 secretion systems (T3SS) and effector proteins [48]. However, some plant

growth promoting rhizobacteria (PGPR) also contain T3SS [49]. Decomposition of macromolecular nitrogen by *Rhizobacter* when other bacteria decrease or cease to do so could imply that these bacteria are beneficial for the plant. However, without isolating this *Rhizobacter* and performing inoculation experiments, it is impossible to determine whether it is a plant pathogen or a potential PGPR.

While the guild appears to contain only a single *Rhizobacter* MAG, the implication of mapping transcriptomic reads at 80% identity is that each MAG represents a taxonomic "cloud" of at least genus-level diversity. A 16S-rRNA survey on the same samples reveals 10 operational taxonomic units (OTUs) of order Burkholderiales which, like *Rhizobacter*, are not assigned a family (Nuccio et al. 2020). Additionally, this guild may include more taxa for which we have no MAGs and therefore were not included in this analysis. Moreover, proteins are a diverse class of molecules, therefore it is reasonable to expect an organism to host multiple types of proteinases. The *Rhizobacter* MAG contains 15 different extracellular proteases

The same *Rhizobacter* MAG (order Burkholderiales) was included in the rhizosphere functional guild based on CAZy expression but had a different expression pattern from other rhizosphere guild members. Similar to its extracellular protease expression pattern, this clade also increased CAZy expression over time in litter-amended rhizosphere and at the last time point in unamended rhizosphere. Recent evidence shows that some Burkholderiales can be primed by root exudates to decompose soil organic carbon [50]. On the flip side, at the last time point the roots begin to age and potentially become a target for degradation, as evidenced by several of the upregulated CAZy groups targeting plant polysaccharides and their downstream byproducts [20]. The upregulation of *Rhizobacter* extracellular protease in later time points, when most of the easily accessible proteins have likely been degraded, may indicate that it has the ability to access N polymers that are less available due to sorption to minerals or occlusion [22].

A recent publication by Nguyen et al. [29] analyzed copy numbers per genome of different structural groups of extracellular proteases, showing that most of these enzymes in soil and aquatic environments are Metallo-, Serine- and Cysteinepeptidases (MP, SP and CP respectively) in decreasing order of abundance. We found these same clusters were dominant in our data, but expression patterns from this study revealed dominance of SP over MP. Most bacterial phyla have dozens of copies of secreted peptidase genes and these genes are enriched especially in Actinobacteria and Proteobacteria, and more so in soil/aquatic environments as opposed to animal associated bacteria. Specifically, betaproteobacteria are enriched in SP (mean 13 SP copies out of a mean of 25 extracellular proteases per genome), with published genomes of isolates of Burkholderiales genera such as *Burkholderia*, *Janthinobacterium* and *Massilia* containing more than 50 and up to 155 copies per genome. As betaproteobacteria are enriched in the *Avena* rhizosphere [25], we might have expected higher expression of SP in this environment compared to bulk soil. Surprisingly, we did not observe that pattern. This could be explained by the high expression of actinobacteria in all groups, as they share the same enrichment pattern with betaproteobacteria (SP>MP).

Conclusions

Extracellular proteases are significantly upregulated by the presence of roots and plant litter. This regulation is taxonomically conserved, and these patterns can be used to define functional guilds. These guilds had some overlap with CAZy defined guilds from the same experiment. They also

exhibited a pattern of temporal succession throughout plant maturation. This succession implies that the change over time in protein decomposers in the rhizosphere could be triggered by the plant through control of exudation or be a response to nitrogen availability. The investigation of metatranscriptomes opened a window to microbial activity that could control plant access to the primary pool of N in soil.

References

1. Curl EA, Truelove B. The Rhizosphere. 1986.

2. Jackson LE, Strauss RB, Firestone MK, Bartolome JW. Plant and soil nitrogen dynamics in California annual grassland. Plant Soil. 1988;110:9–17.

3. Vitousek P, Howarth R. Nitrogen limitation on land and in the sea: How can it occur? Biogeochemistry. 1991;13. doi:10.1007/bf00002772.

4. MYROLD, D D. Transformations of nitrogen. Principles and applications of soil microbiology. 1999;550. https://ci.nii.ac.jp/naid/10011797564/. Accessed 3 Apr 2020.

5. Jones DL, Healey JR, Willett VB, Farrar JF, Hodge A. Dissolved organic nitrogen uptake by plants—an important N uptake pathway? Soil Biol Biochem. 2005;37:413–23.

6. Braun J, Mooshammer M, Wanek W, Prommer J, Walker TWN, Rütting T, et al. Full 15N tracer accounting to revisit major assumptions of 15N isotope pool dilution approaches for gross nitrogen mineralization. Soil Biol Biochem. 2018;117:16–26.

7. Moreau D, Bardgett RD, Finlay RD, Jones DL, Philippot L. A plant perspective on nitrogen cycling in the rhizosphere. Funct Ecol. 2019;33:540–52.

8. Harte J, Kinzig AP. Mutualism and competition between plants and decomposers: implications for nutrient allocation in ecosystems. Am Nat. 1993;141:829–46.

9. Paul EA, Clark FE. Soil microbiology and biochemistry. --Academic Press. New York, USA. 1996.

10. Kinzig AP, Harte J. Selection of micro-organisms in a spatially explicit environment and implications for plant access to nitrogen. Journal of Ecology. 1998;86:841–53. doi:10.1046/j.1365-2745.1998.8650841.x.

11. Dijkstra FA, Carrillo Y, Pendall E, Morgan JA. Rhizosphere priming: a nutrient perspective. Front Microbiol. 2013;4:216.

12. Zhu B, Gutknecht JLM, Herman DJ, Keck DC, Firestone MK, Cheng W. Rhizosphere priming effects on soil carbon and nitrogen mineralization. Soil Biol Biochem. 2014;76:183–92.

13. Meier IC, Finzi AC, Phillips RP. Root exudates increase N availability by stimulating microbial turnover of fast-cycling N pools. Soil Biol Biochem. 2017;106:119–28.

14. Zhalnina K, Louie KB, Hao Z, Mansoori N, da Rocha UN, Shi S, et al. Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. Nat Microbiol. 2018;3:470–80.

15. Chapin FS, Stuart Chapin F, Matson PA, Mooney HA. Principles of Terrestrial Ecosystem Ecology.

2002. doi:10.1007/b97397.

16. Schimel JP, Bennett J. NITROGEN MINERALIZATION: CHALLENGES OF A CHANGING PARADIGM. Ecology. 2004;85:591–602. doi:10.1890/03-8002.

17. Pett-Ridge J, Firestone MK. Using stable isotopes to explore root-microbe-mineral interactions in soil. Rhizosphere. 2017;3:244–53.

18. Jaeger CH 3rd, Lindow SE, Miller W, Clark E, Firestone MK. Mapping of sugar and amino acid availability in soil around roots with bacterial sensors of sucrose and tryptophan. Appl Environ Microbiol. 1999;65:2685–90.

19. Herman DJ, Johnson KK, Jaeger CH, Schwartz E, Firestone MK. Root Influence on Nitrogen Mineralization and Nitrification in Avena barbata Rhizosphere Soil. Soil Sci Soc Am J. 2006;70:1504–11.

20. Nuccio EE, Starr E, Karaoz U, Brodie EL, Zhou J, Tringe SG, et al. Niche differentiation is spatially and temporally regulated in the rhizosphere. ISME J. 2020. doi:10.1038/s41396-019-0582-x.

21. Nguyen T, Kleber M, Myrold DD. Contribution of different catalytic types of peptidases to soil proteolytic activity. Soil Biol Biochem. 2019;138:107578.

22. Fukumasu J, Shaw LJ. The role of macro-aggregation in regulating enzymatic depolymerization of soil organic nitrogen. Soil Biol Biochem. 2017;115:100–8.

23. Greenfield LM, Hill PW, Seaton FM, Paterson E, Baggs EM, Jones DL. Is soluble protein mineralisation and protease activity in soil regulated by supply or demand? Soil Biol Biochem. 2020;150:108007.

24. DeAngelis KM, Lindow SE, Firestone MK. Bacterial quorum sensing and nitrogen cycling in rhizosphere soil. FEMS Microbiol Ecol. 2008;66:197–207.

25. Shi S, Nuccio E, Herman DJ, Rijkers R, Estera K, Li J, et al. Successional Trajectories of Rhizosphere Bacterial Communities over Consecutive Seasons. MBio. 2015;6:e00746.

26. Huang Y, Niu B, Gao Y, Fu L, Li W. CD-HIT Suite: a web server for clustering and comparing biological sequences. Bioinformatics. 2010;26:680–2.

27. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119.

28. Rawlings ND, Barrett AJ, Finn R. Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. 2016;44:D343–50.

29. Nguyen TTH, Myrold DD, Mueller RS. Distributions of Extracellular Peptidases Across Prokaryotic Genomes Reflect Phylogeny and Habitat. Front Microbiol. 2019;10:413.

30. Bushnell B. BBTools software package. URL http://sourceforge net/projects/bbmap. 2014.

31. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Liaw WHA, Lumley T, et al. gplots: Various R programming tools for plotting data. 2015. https://www.scienceopen.com/document?vid=0e5d8e31-1fe4-492f-a3d8-8cd71b2b8ad9.

32. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer; 2016.

33. Starr EP, Shi S, Blazewicz SJ, Probst AJ, Herman DJ, Firestone MK, et al. Stable isotope informed genome-resolved metagenomics reveals that Saccharibacteria utilize microbially-processed plant-derived carbon. Microbiome. 2018;6:122.

34. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27–30.

35. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res. 2019;47:e47.

36. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.

37. Oksanen J, Kindt R, Legendre P, O'Hara B, Simpson GL, Solymos P, et al. The vegan package. Community ecology package [http://r-forge r-project org/projects/vegan/]. 2008. https://www.researchgate.net/profile/Gavin_Simpson/publication/228339454_The_vegan_Package/links/0912f50be86bc29a7f000000/The-vegan-Package.pdf.

38. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics. 2019. doi:10.1093/bioinformatics/btz848.

39. Lee MD. GToTree: a user-friendly workflow for phylogenomics. Bioinformatics. 2019;35:4162–4.

40. Bai Y, Müller DB, Srinivas G, Garrido-Oter R, Potthoff E, Rott M, et al. Functional overlap of the Arabidopsis leaf and root microbiota. Nature. 2015;528:364–9.

41. Olk DC. Organic Forms of Soil Nitrogen. Nitrogen in Agricultural Systems. 2008;:57–100.

42. Cheng W, Parton WJ, Gonzalez-Meler MA, Phillips R, Asao S, McNickle GG, et al. Synthesis and modeling perspectives of rhizosphere priming. New Phytol. 2014;201:31–44.

43. Allison SD, Vitousek PM. Responses of extracellular enzymes to simple and complex nutrient inputs. Soil Biology and Biochemistry. 2005;37:937–44. doi:10.1016/j.soilbio.2004.09.014.

44. Noll L, Zhang S, Zheng Q, Hu Y, Wanek W. Wide-spread limitation of soil organic nitrogen transformations by substrate availability and not by extracellular enzyme content. Soil Biol Biochem. 2019;133:37–49.

45. Geisseler D, Horwath WR. Relationship between carbon and nitrogen availability and extracellular enzyme activities in soil. Pedobiologia . 2009;53:87–98.

46. Berg G, Eberl L, Hartmann A. The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. Environmental Microbiology. 2005;7:1673–85. doi:10.1111/j.1462-2920.2005.00891.x.

47. DeAngelis KM, Brodie EL, DeSantis TZ, Andersen GL, Lindow SE, Firestone MK. Selective progressive response of soil microbial community to wild oat roots. ISME J. 2009;3:168–78.

48. Starr EP, Shi S, Blazewicz SJ, Koch BJ, Probst AJ, Hungate BA, et al. Stable isotope informed genome-resolved metagenomics uncovers potential trophic interactions in rhizosphere soil. Cold Spring Harbor Laboratory. 2020;:2020.08.21.262063. doi:10.1101/2020.08.21.262063.

49. Dias GM, de Sousa Pires A, Grilo VS, Castro MR, de Figueiredo Vilela L, Neves BC. Comparative genomics of Paraburkholderia kururiensis and its potential in bioremediation, biofertilization, and biocontrol of plant pathogens. Microbiologyopen. 2019;8:e00801.

50. Wilhelm RC, DeRito CM, Shapleigh JP, Madsen EL, Buckley DH. Phenolic acid-degrading Paraburkholderia prime decomposition in forest soil. Cold Spring Harbor Laboratory. 2020;:2020.09.28.317347. doi:10.1101/2020.09.28.317347.

Declarations

Ethics approval and consent to participate: not applicable
Consent for publication: not applicable
Competing interests: The authors declare that they have no competing interests

Availability of data and material: All data used in this publication, including raw reads and genome collection, has been published previously (Nuccio et al., 2020, see sup. table S1 within).

Acknowledgements

The authors would like to thank Katerina Estera-Molina, Rachel Hestrin, Rachel Neurath, Evan Starr and Nameer Baker for their support and helpful feedback on the analyses.

Funding

This research was supported by the U.S. Department of Energy Office of Science, Office of Biological and Environmental Research Genomic Science program under Awards DE-SC0020163 and DE-SC0016247 to M.K.F at UC Berkeley and awards SCW1589 and SCW1678 to J.P-R. at Lawrence Livermore National Laboratory. Sequencing was conducted as part of Community Sequencing Awards 1487 to JPR and 1472 to MKF. Work conducted at LLNL was contributed under the auspices of the US Department of Energy under Contract DE-AC52-07NA27344.

Authors' contributions: ETS performed the data analysis, data interpretation and drafted the manuscript. EEN conceived and designed the study, interpreted the data and revised the manuscript. JPR interpreted the data and substantially revised the manuscript. MKF conceived and designed the work, interpreted the data and substantially revised the manuscript. All authors approved the submitted version and agreed both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.