Expression of macromolecular organic nitrogen degrading enzymes identifies potential mediators of soil organic N availability to an annual grass

Running title: Microbial protease expression in rhizosphere

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

[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

Email addresses:

[#]Address correspondence to: Mary Firestone mkfstone@berkeley.edu or Ella Sieradzki ella.shir@gmail.com

The authors declare no competing financial interests.

Abstract

Nitrogen (N) commonly limits terrestrial plant growth partly because most soil-N is present as macromolecular organic compounds and not directly available to plants. Soil microbes degrade these large N-containing substrates to gradually release plant-available inorganic-N throughout the growing season, potentially meeting plant demand. Knowing which microbes are responsible for release of organic N, as well as their spatiotemporal patterns of activity, can enable microbial management strategies that increase plant access to soil-N and reduce dependency on fertilizer-N. We used time-resolved metatranscriptomes to follow taxonomy-resolved differential expression of N-depolymerizing enzymes.

Taxonomic groups show adaptations based on extracellular proteases to specialized habitats in soil characterized by presence (Betaproteobacteria) or absence (Thaumarcheota) of live roots and root detritus (Deltaproteobacteria and Fungi). A similar increase of eukaryotic chitinases near root detritus hints at predation of fungi. Others demonstrate temporal patterns such as increasing expression over time, implying increased competitiveness with substrate depletion (Chloroflexi). Phylotypes from the same genus can have different potential benefits to the plant based on protease expression (e.g., *Janthinobacterium*), which should be considered when designing bioaugmentation. We identify one *Janthinobacterium* phylotype and two Burkholderiales that may be candidates for bioaugmentation near young roots and a *Rhizobacter* which could benefit mature roots.

Keywords

Rhizosphere, microbiome, protease, chitinase, functional guilds, metatranscriptome, gene expression, litter

41 **Introduction**

42 Plants are commonly limited by nitrogen (N) in temperate soils since their access to the
43 largest soil N pools is constrained by the activity of microbes responsible for ammonification of
44 macromolecular nitrogen and N_2 fixation [1–3]. To a large degree, plants depend on microbial
45 degradation of macromolecular forms of N such as proteins and chitin [4–6] and can potentially
46 increase N availability by releasing exudates that stimulate microbial turnover of organic N pools
47 [7–10]. Depolymerization of high molecular weight detrital organic N is a primary rate-limiting
48 step in soil N mineralization [11, 12] and depends on the activity of extracellular enzymes such
49 as lysozyme, protease, chitinase, nuclease and urease. The resulting N monomers are taken up by
50 microorganisms, and roughly 30% of the amino-acid carbon is respired, leading to excretion of
51 excess ammonium which can benefit the plant [13–16]. While the bulk-scale activity of
52 extracellular N degrading enzymes in soil is established, the spatial and temporal dynamics of
53 gene expression underpinning N mineralization are unknown. Additionally, soil microorganisms
54 with the genomic capacity for organic N degradation are diverse [17], and it is unclear whether
55 there is niche partitioning amongst them [18].

56 Patterns and controls of soil N depolymerization and mineralization may be highly
57 dependent on the habitat (e.g., rhizosphere, detritusphere), since availability of organic-N
58 substrates varies in both time and space, as does the prevalence of fungal, faunal, and bacterial
59 degraders, which can enable or limit N-cycling. In the rhizosphere, organic N is available as
60 amino acids, nucleotides, niacin, and choline (derived from plant exudates [10]), lignoproteins
61 and aromatics (from sloughed off cells), nucleic acids, and microbial cell wall amino sugar
62 polymers (N-acetylglucosamine, N-acetylmuramic acid) derived from the bloom of cells that
63 develops as roots grow [18]. Prior work indicates that root exudates can increase degradation of

64 soil organic matter by up to 380% [19]. In contrast, the detritusphere has a relatively higher
65 proportion of aromatic N and lignoproteins. Bulk protease enzyme activity has been shown to be
66 enhanced by both litter addition [20] and root exudates [21].

67 Soil N pools and by proxy, N mineralization activity, change with time [22] in part due to
68 the succession of both exudate quality and microbial communities [10, 23]. At the bulk soil level,
69 activity of extracellular proteases is thought to drive soil N cycling [12], but what controls this
70 activity is not always clear [24]. Subtle distinctions caused by time and soil habitat are likely
71 masked because most studies are conducted on whole soils, which contain a mixture of
72 rhizosphere, detritusphere and bulk regions. In grassland soils, the rhizosphere is a particularly
73 critical hotspot for microbial biomass and activity due to rapid recycling of root exudation and
74 debris, and functions as a quasi-digestive system decomposing molecules inaccessible to plants.
75 However, root exudation, which serves as a carbon source for rhizosphere bacteria, declines near
76 older root sections [25]. Therefore, it has been hypothesized that bacteria in a mature rhizosphere
77 environment may be forced to target less labile, higher C:N sources of organic N such as plant
78 litter [26]. Indeed, it has been shown that the mature *Avena* sp. rhizosphere has higher rates of
79 gross nitrogen mineralization and ammonia consumption compared to the young rhizosphere
80 [28].

81 To determine spatial and temporal patterns and identify which groups of microbes were
82 primarily responsible for decomposition of protein and chitin, we assessed the expression of
83 genes that code for enzymes degrading macromolecular organic N near actively growing and
84 decomposing roots of a common annual grass *Avena fatua* (wild oat grass). We used wild oat as
85 a potential source for putative beneficial microbes to domestic oat, as wild oat in our system does
86 not experience fertilization and therefore needs to acquire N from soil organic matter via its

87 microbiome. We analyzed 48 metatranscriptomes collected over three weeks of active root
88 growth, using a well-developed plant mesocosm approach[23, 26–28] with a fully-characterized
89 annual grassland soil from Northern California. A different subset of these metatranscriptomes
90 was previously used to analyze expression of genes coding for degradation of carbohydrates, in
91 an effort to determine the effect of roots and litter on carbon cycling [28]. The previous study
92 provided important context that allowed us to compare trends in carbon and nitrogen cycling.
93 The primary goal of the analysis presented here was to discover the roles of indigenous
94 rhizosphere microorganisms in macromolecular nitrogen degradation by extracellular enzymes
95 throughout root growth and aging. This identification is the first step towards *in-situ* inoculation.
96 Inoculation of seeds with plant growth promoting bacteria has shown promise in reducing the
97 need for fertilizers[29] and improving plant biomass[30] as well as rhizosphere available
98 nitrogen[31].
99

100 **Materials and methods**

101 *Experimental design, sample collection and sequencing*

102 The experimental design, sample collection and sequence data processing are described
103 in detail in [28]. Briefly, common wild oat *Avena fatua* was grown for six weeks in rhizobox
104 microcosms containing soil from the Hopland Research and Extension Center (HREC) in
105 northern California, a Bearwallow–Hellman loam (pH 5.6, 2% total C, include 7th approx. name
106 here) packed at field bulk density. Roots were then allowed to grow into a sidecar soil region
107 with a transparent wall, where the root growth timeline was marked at 3 days, 6 days, 12 days
108 and 22 days. In half of the sidecars, the soil was amended with dried *A. fatua* root detritus (C:N =
109 13.4) chopped to 1 mm. Bulk soil bags, inaccessible to roots, were placed in each sidecar; a bulk

110 soil treatment amended with root detritus was also included. At each timepoint, three replicate
111 microcosms were destructively harvested for paired rhizosphere and bulk soil, both detritus
112 amended and unamended, yielding a total of 48 samples representing four treatments at four
113 timepoints: rhizosphere, rhizosphere + litter, bulk, and bulk + litter (sup. fig. S1).
114 Harvested soil (1 g) was placed immediately in Lifeguard Soil Preservation Reagent
115 (MoBio) and processed according to the company protocol. Roots and supernatant were
116 removed, and the soil was stored in -80°C. DNA and RNA were co-extracted using a phenol-
117 chloroform procedure [32, 33] and separated with an AllPrep kit (Qiagen). RNA was DNase
118 treated (TURBO DNase, Thermo-Fisher Scientific), depleted in ribosomal RNA (Ribo-Zero
119 rRNA Removal Kit, Epicentre) and reverse transcribed into cDNA. Metatranscriptomes were
120 sequenced for 48 samples on an Illumina HiSeq 2000 2x150 (TruSeq SBS v3) at the Joint
121 Genome Institute (JGI), in Walnut Creek CA.
122
123 *Expression of macromolecular N degrading genes identified in denovo assembled*
124 *metatranscriptomes*
125 Raw reads were quality-trimmed (Q20) and rRNA and tRNA reads were removed.
126 Library size, evenness, richness and Shannon diversity were comparable between experimental
127 groups with a mean library size of 43 M paired end reads after filtering. In contrast to the single
128 approach used by Nuccio et al. [28], we also *denovo* assembled quality-controlled
129 metatranscriptomic reads into contigs within each sample. Contigs smaller than 200bp were
130 discarded and the remaining contigs from all samples were clustered at 99% identity with cd-hit-
131 est keeping the longest sequence as the cluster representative [34]. Open reading frames (ORFs)
132 were predicted using Prodigal [35]. Extracellular protease ORFs were identified by reciprocal

133 BLAST to extracellular peptidases from the MEROPS database [17, 36]. We acknowledge that
134 this method may generate a conservative estimate of the number of proteases detected. ORFs
135 were assigned a peptidase group (serine-, metallo-, cysteine-peptidase and others) by their best
136 reciprocal BLAST hit. Taxonomy was determined by best BLAST hit to the NCBI non-
137 redundant database (NR, access date July 29th 2019). Chitinase ORFs were identified using six
138 chitinase hidden Markov models (HMMs) from the Kyoto Encyclopedia of Genes and Genomes
139 (KEGG): K01183, K03791, K13381, K17523, K17524 and K17525. Only the first three were
140 detected in our dataset. Lysozyme ORFs were identified using the HMM for KEGG orthology
141 K07273, extracellular nuclease with K01150 (codes for Dns gene, undetected) and K07004, and
142 urease subunit A, B and C with K01428, K01429 and K01430 respectively. Reads were then
143 mapped back to ORFs requiring minimum identity 95% and 75% breadth using bbmap [37].
144 Read counts were normalized using DESeq2 [38]. Heat maps of normalized counts were
145 generated in R using the heatmap2 function in gplots [39]. Normalized transcript counts per gene
146 per time point were compared between groups (rhizosphere, litter and litter-amended
147 rhizosphere) using ANOVA and Tukey post-hoc test. Boxplots were generated in ggplot2 [40].
148
149 *Expression of extracellular protease genes from a curated collection of Hopland-soil genomes*
150 The quality-controlled reads were mapped using BBsplit [37] requiring 80% identity
151 against a dereplicated reference database of 282 HREC soil genomes including isolates [10],
152 single amplified genomes (SAGs) [28], metagenomic assembled genomes (MAGs) (NCBI
153 PRJNA517182) and stable isotope probing MAGs (SIP-MAGs) [41]. On average, 12.3% (range
154 6.2-31.9%) of the reads per library mapped unambiguously to genomes from the reference
155 database. This additional approach was taken to investigate gene expression within the context of

156 a genome and to search for overlap in guild membership between extracellular protease defined
157 guilds and previously defined CAZy guilds [28]. Three MAGs that were classified as unknown
158 or domain bacteria in previous studies were assigned taxonomy using GTDB-Tk version 1.3.0
159 [42, 43]. Verification of the taxonomic assignment of MAG Burkholderiales_62_29 was done
160 using GToTree [44] with complete reference genomes of Betaproteobacteria from RefSeq (Feb
161 28, 2020).

162 Open reading frames (ORFs) were predicted in all genomes using prodigal [35] and
163 annotated using KEGG [45] and ggKbase (http://ggkbase.berkeley.edu). Extracellular proteases,
164 which do not have hidden Markov models (HMMs) capable of separating them from intracellular
165 proteases, were identified by gene nucleotide identity of at least 90% and coverage of at least
166 60% to *de novo* assembled ORFs of extracellular protease from the metatranscriptomes. Gene
167 counts were identified using Rsubread featureCounts [46].

168 Differential expression of chitinases in the genome collection was already performed in a
169 previous publication on expression of carbohydrate active enzymes (CAZy) and was therefore
170 not repeated here [28].

171

172 *Statistical analyses*

173 All features and their abundance (represented by metatranscriptomic read counts
174 normalized to sequencing depth) were analyzed with DESeq2 [38] requiring p-value < 0.05
175 (adjusted for multiple comparisons). Ordination and visualization were conducted in R using
176 ggplot2 [40] and vegan [47]; PERMANOVA (vegan function adonis) was used to detect
177 significant treatment factors affecting expression of nitrogen depolymerization genes (protease
178 expression ~ location * treatment * time). We define 'guilds' as groups of taxa with similar gene

179 upregulation patterns of extracellular proteases in both time and soil habitat. Guilds were
180 assigned by hierarchical clustering based on differential expression of extracellular proteases
181 compared to unamended bulk soil, generating four functional guilds: “Rhizosphere”,
182 “Detritusphere”, “Aging root”, and “Low response". Effects of the four experimental treatments
183 (bulk, rhizosphere, litter and litter-amended rhizosphere) on protease gene expression were
184 assessed by ANOVA (p adjusted for multiple comparisons), N=3 per time point.
185
186 *Data availability*
187 The R code used in this work is publicly available at
188 https://github.com/ellasiera/Protease_ISME_2022. All data used in this publication, including
189 raw reads and genome collection, is publicly available. Metagenomes assembled MAGs can be
190 found in NCBI PRJNA517182, stable isotope probing MAGs in http://ggkbase.berkeley.edu/,
191 single amplified genomes in IMG under study name Mediterranean Grassland Soil Metagenome
192 and single amplified genomes in IMG, see sup. table S2 in Nuccio et al., 2020 for accession
193 numbers [28]. Raw reads can be found in JGI IMG. For JGI ID’s (accession numbers) see
194 supplemental table S1 in Nuccio et al. 2020 [28].
195
196 **Results**
197 *Expression of enzymes targeting macromolecular N*
198 Extracellular enzymes depolymerizing macromolecular N in soil mainly target proteins, cell wall
199 components, and nucleic acids[48]. Multiple genes for these functions were identified in this
200 analysis: extracellular nuclease (Xds), urease (UreABC), lysozyme targeting peptidoglycan,
201 chitinase (*chit1*) targeting fungal cell walls and insect exoskeleton and extracellular proteases.

202 However, expression of extracellular proteases was an order of magnitude higher than the other
203 extracellular N degrading enzymes (sup. fig. S2), and like chitinases, was affected by the
204 presence of litter and roots. In contrast, extracellular nuclease (Xds), urease (UreABC) and
205 lysozyme were not influenced by either living or decaying roots, thus, we chose to focus on
206 extracellular proteases and chitinases.

207 Normalized transcript abundance for 4948 extracellular protease genes from bacteria
208 (4846) and fungi (102) was significantly affected by time (3-way PERMANOVA, $F=2.8$,
209 $p=0.038$), litter amendment ($F=118.9$, $p<0.001$) as well as interactions between time:treatment
210 ($F=26.9$, $p<0.001$), time:location ($F=7.7$, $p<0.001$) and location:treatment ($F=15.5$, $p<0.001$).
211 Expression in unamended soil (no litter) generally increased slightly over time (fig. 1A), whereas
212 litter-amended soil had initially high expression that then decreased and leveled off over time
213 (fig. 1B).

214 The 73 distinct chitinase transcripts were expressed at a substantially lower level than the
215 extracellular proteases. The expression of the *chit1* gene increased over time in litter-amended
216 soils and was higher than in the unamended treatments (3-way PERMANOVA, $p<0.001$) (fig.
217 1C). Expression of fungal chitinases was lower than bacterial chitinases, and both were generally
218 lower than eukaryotic chitinases (fig. 1D). Transcripts for chitinase KEGG orthologs *CHI3L1_2*,
219 *CHI3L1_4* and *chiA* were not detected at all and the putative chitinase K03791 had low
220 expression and no significant effects with time or treatment (data not shown). Chitinase CHID1
221 was detected but was most closely related to plants and therefore disregarded. A companion
222 study analyzed differential expression of chitinases within our curated genome collection and
223 while only one MAG was found to upregulate chitinase, it did so in the presence of litter [28].
224

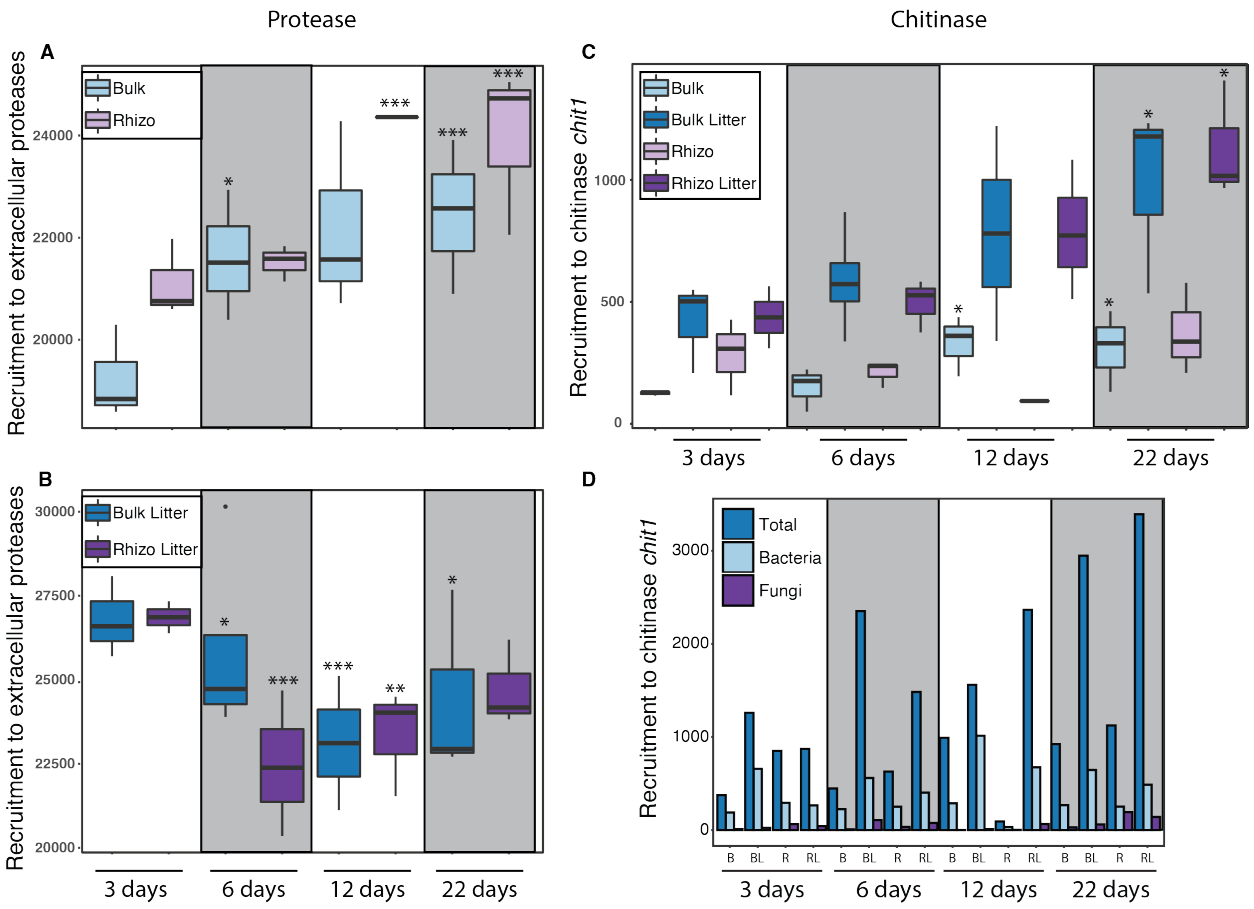


Figure 1: Aggregated expression of extracellular proteases in bulk and rhizosphere soils from common wild oat *Avena fatua* microcosms, grown without root litter amendment (A), or with litter amendment (B); expression of chitinase gene *chit1* (C) and mean counts of chitinases at the domain level (D) for unamended bulk soil (B), litter-amended bulk soil (BL), unamended rhizosphere (R) and litter-amended rhizosphere (R L). 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. Statistical significance of linear regression compared to 3 days within the same location and treatment group: * p value <0.05, ** p value <0.01, *** p value < 0.001.

Structural groups of extracellular proteases

237 Structural groups of extracellular proteases can be soil-specific and pH dependent [49].
238 The main groups of extracellular proteases found in our assembled metatranscriptomes were
239 serine-, metallo- and cysteine-proteases (2679, 1949 and 496, respectively out of 5295 variants
240 clustered at 99% identity). Aggregated expression patterns of all variants in each group reflected
241 the same order (serine>metallo>cysteine). Expression of serine- was consistently higher than
242 metallo-protease across all treatments (ANOVA $F=2392$, $p=0$) (fig. 2, sup. table S1). Expression
243 of serine-proteases was also significantly greater in the presence of litter compared to no litter
244 (diff=2337, $p<0.001$), but root litter amendment did not affect expression of metallo- or cysteine-
245 proteases, and there was no significant effect of time or location on these structural groups (sup.
246 table S1).
247

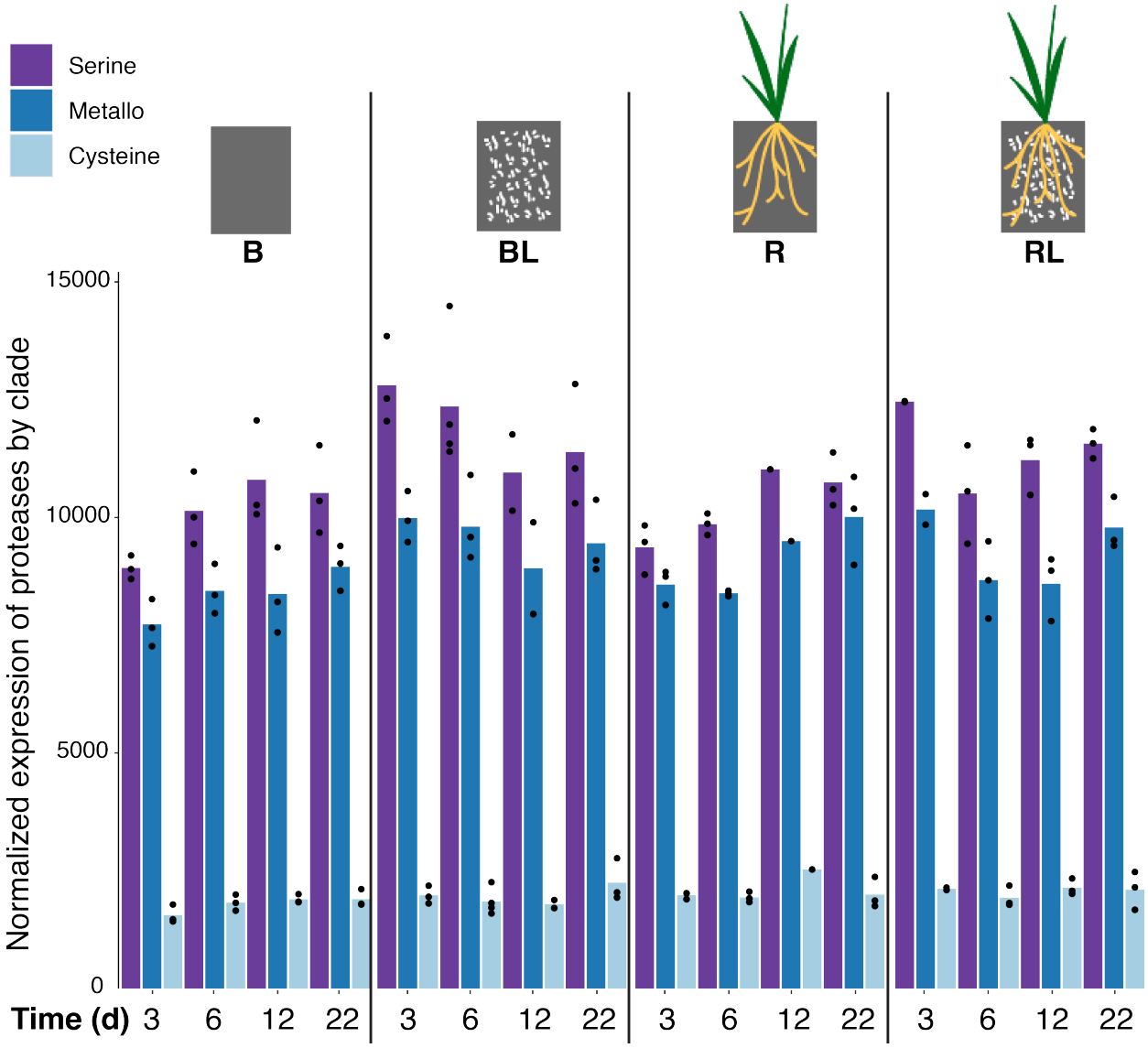


Figure 2: Aggregated normalized expression of extracellular protease structural groups over time in bulk, rhizosphere and root litter-amended soils. Normalized expression per replicate is plotted as dots on top of the bars. Experimental groups shown per time point are (left to right): unamended bulk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

256 *Taxonomy of extracellular proteases*

257 *De novo* assembled ORFs of extracellular proteases were taxonomically assigned by
258 BLASTP best hit against the NCBI non-redundant database (NR). Since extracellular proteases
259 are not marker genes, we considered taxonomic assignments only at the order level or higher,
260 with the exception of *Rhizobacter*, for which amino acid percent identity values were extremely
261 high (84±8.4% amino acid identity).

262 In the rhizosphere, Betaproteobacteria proteases were significantly upregulated
263 (ANOVA, $p<0.05$) (fig. 3A). Of 547 variants of proteases from this class, 442 were assigned as
264 Burkholderiales. In contrast, at most timepoints, proteases of Cyanobacteria and Thaumarchaeota
265 were significantly upregulated in bulk soil compared to the rhizosphere (fig. 3B; sup. table S2).
266 In multiple other taxonomic groups, proteases were significantly upregulated only in the
267 presence of litter: Fungi (fig. 3C), class Deltaproteobacteria (fig. 3D), as well as highly
268 represented phyla Bacterioidetes and Verrucomicrobia (sup. fig. S2; sup. table S2) and classes
269 Chitinophagia and Gammaproteobacteria (sup. table S3). Interestingly, protease expression
270 declined with time for the predatory bacteria Myxococcales (fig. 3E), Bdellovibrionales and
271 Cytophagia, while clades such as Chloroflexi (fig. 3F) and Actinobacteria had increased protease
272 expression at the final sampling point. While the Actinobacteria and Acidobacteria had a high
273 number of variants and high expression of proteases, we did not detect a significant effect of
274 either time or soil habitat/ amendment. Normalized protease expression data are summarized by
275 phylum (sup. fig. S3), class (sup. fig. S4) and order (sup. fig. S5), and ANOVA F and p values
276 are in sup. table S2, sup. table S3 and sup. table S4 respectively.

277

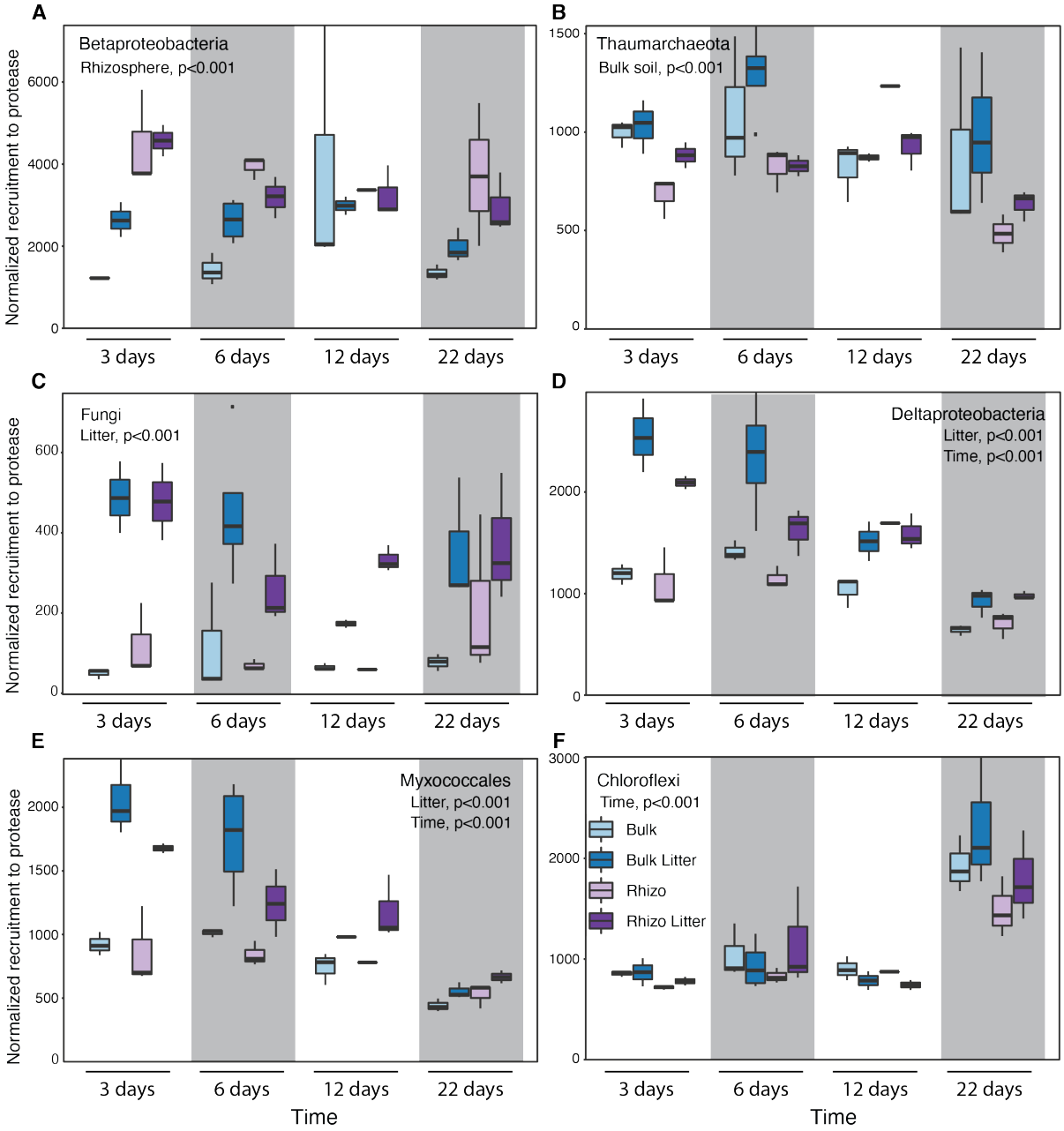


Figure 3: Expression of extracellular proteases over time for select taxonomic groups: (A) Betaprotebacteria (n=547) were upregulated in the rhizosphere, (B) Thaumarchaeota (n=112) were downregulated in the rhizosphere, (C) Fungi (n=99) were upregulated in the presence of litter, (D) Detaproteobacteria (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.

285 Note that as expression levels varied over taxonomic groups and to emphasize patterns, the y
286 axes do not have the same scale.
287
288 *Functional guilds*
289 To define functional guilds with a population-centric analysis, we mapped transcriptome reads to
290 a collection of 282 genomes and metagenome-assembled genomes generated from the same soil
291 and site[28]. Read counts were used to determine differential expression compared to
292 unamended bulk soil. Each genome contained multiple genes coding for extracellular proteases.
293 Hierarchical clustering of the mean differential expression of extracellular protease genes within
294 each genome was used to define four functional guilds: "rhizosphere", "detritusphere", "aging
295 root" and "low response" (fig. 4). These guilds include bacteria that upregulate expression of
296 extracellular proteases in the presence of live roots (rhizosphere), dead roots (detritusphere), live
297 roots that are beginning to senesce (aging root) or at low but significant levels without a specific
298 pattern (low response)[28]. Similarly, hierarchical clustering based on differential expression of
299 the most highly upregulated *de novo* assembled ORFs compared to unamended bulk soil
300 revealed clear rhizosphere and detritusphere guilds (sup. fig. S6). There was some overlap
301 between guilds identified here and guilds previously identified based on expression of CAZy
302 genes[28], but the degree of overlap varied by guild (sup. fig. S7).
303 Within the aging root guild, we noticed one member, representing an aggregated
304 population transcriptome, that increased differential expression of extracellular proteases at the
305 last time point in the presence of both roots and litter more than would have been predicted by
306 the sum of the treatments alone. As this pattern differed from the rest of the guild, we identified
307 this member as a sub-guild labeled "late synergist" due to its expression pattern. This MAG,

308 Burkholderiales_62_29, was identified by two independent phylogenomic analyses as

309 *Rhizobacter*. Burkholderiales_62_29 has 16 different extracellular protease genes, all of which

310 had similar significant upregulation patterns. While the sub-guild here contained only a single

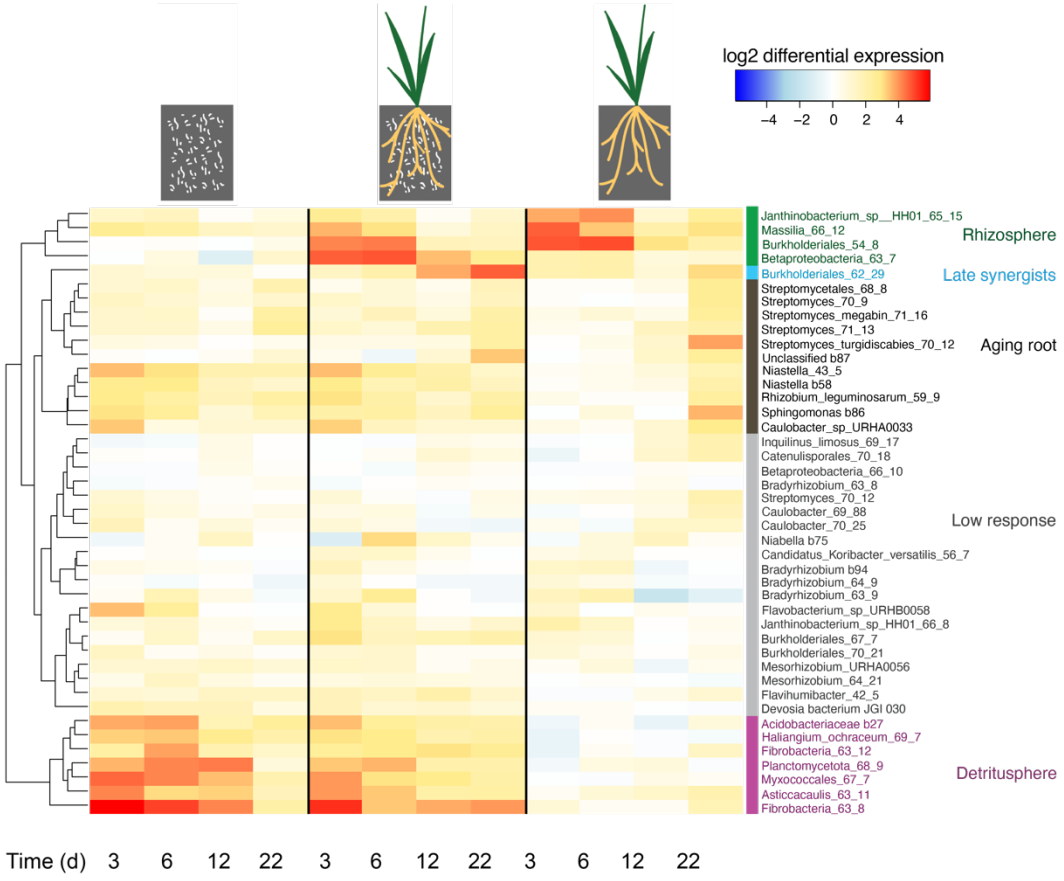
311 *Rhizobacter* MAG, the implication of mapping transcriptomic reads at 80% identity is that each

312 MAG represents a taxonomic "cloud" of at least genus-level diversity. A 16S-rRNA survey of

313 the same samples revealed 10 operational taxonomic units (OTUs) of order Burkholderiales

314 which, like *Rhizobacter*, are not assigned a family [28]. Additionally, this sub-guild may include

315 more taxa for which we have no MAGs and therefore could not be included in this analysis.



316

317 **Figure 4**: Functional guilds defined by hierarchical clustering of extracellular protease

318 differential gene expression during a 22-day *Avena fatua* microcosm experiment. Each row

319 represents the mean differential expression of extracellular protease genes mapped to a reference

genome. Treatments are (left to right): litter-amended bulk soil, litter-amended rhizosphere and unamended rhizosphere. Time points are indicated at the bottom in days. We note that a genome may contain more than one protease gene and that reads were mapped at 80% identity, 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.

Discussion

Our current understanding of macromolecular nitrogen depolymerization by microbes in soil is based almost entirely on measurement of bulk mineralization rates and genomic surveys. Linking rates and genetic potential is challenging, as rates cannot be broken down by taxonomy and gene presence represents a blueprint which may or may not be acted upon. For example, we show here that of the organic N-degrading enzymes detected in this study, only protease and chitinase are differentially expressed near roots. However, this change in expression was not reflected in changes in the overall microbial community composition [28]. Based on genomic potentials previously reported by Nguyen et al.[17], we hypothesized that expression of structural groups of proteases would be dominated by metalloproteases[17] but expression of serineproteases proved to be significantly higher across experimental groups in this study. Therefore, to identify bacterial candidates that may be beneficial to the plant by improving nitrogen accessibility, analysis of transcripts which provide both taxonomic and metabolic information is crucial.

Chitinase expression, while two orders of magnitude lower than that of proteases, was significantly higher and increased over time in litter-amended samples. This effect could imply that the litter attracted saprotrophic fungi and possibly arthropods, the cell walls of which contain

343 chitin [50]. Indeed, expression of fungal proteases was significantly higher in the presence of
344 litter. As most of the chitinases were most closely related to eukaryotic enzymes, it is likely that
345 the chitin-containing hyphae were preyed upon by soil fauna.

346 We found that in unamended soil, expression of extracellular proteases was higher in the
347 rhizosphere compared to bulk soil, potentially resulting from input of carbon from root exudates
348 creating a higher demand for nitrogen as well as microbial competition over inorganic nitrogen
349 with the plant. Similarly, DeAngelis et al. showed that protease specific activity was significantly
350 higher in the young rhizosphere compared to bulk soil, but activity (based on enzyme assays)
351 was not significantly different based on root age [27]. In litter-amended soil, expression of
352 extracellular protease was highest regardless of root presence at 3 days, suggesting that litter-
353 added carbon overwhelms the effect of root exudates at this early point in time. At 12 and 22
354 days there was no difference between litter-amended and unamended soil in either rhizosphere or
355 bulk soil, suggesting that the effect of litter had waned, likely due to substrate degradation. In
356 support, expression of carbohydrate active enzymes after litter addition generally decreased over
357 time [28].

358 Comparing membership between guilds previously defined by CAZy expression [28] and
359 guilds defined here by hierarchical clustering of extracellular protease expression revealed guild-
360 dependent trends. In the rhizosphere guilds, comprised primarily of Betaproteobacteria known to
361 be commonly associated with rhizosphere soil[23, 51, 52], we found high similarity in
362 membership, whereas in the detritusphere guilds there was low similarity. This implies that in the
363 rhizosphere, the same bacteria break down complex carbon and complex nitrogen, whereas in
364 bulk soil these processes are performed by different taxa. The functional guilds identified here
365 exhibited a pattern of temporal succession throughout plant maturation, where expression of

366 protein degrading enzyme genes in the rhizosphere could be influenced by the host plant through
367 control of exudation, or alternatively, may be a response to competition with the plant for labile
368 nitrogen. Moreover, Nuccio et al. [28], analyzing a different subset of this experimental data,
369 showed that microbial community composition (by 16S-rRNA) in each experimental group
370 changed very little compared to gene expression over the course of this experiment [28].
371 Therefore, the changing expression patterns over time that we observed are likely related to
372 changes in environmental conditions or cues, such as macromolecular N availability and
373 inorganic N availability, as opposed to wholesale shifts in community composition.
374 Augmentation of soil or seeds with bacteria that specialize in macromolecular organic N
375 degradation could be used to reduce the use of fertilizers. Ideally, bioaugmentation should use
376 bacteria endemic to the specific plant and soil in order to provide strong competitors[53].
377 Different patterns of expression of extracellular proteases by various taxonomic groups indicate
378 that potential beneficial partners for the plant change throughout root growth. Hence, functional
379 guild characterization can be useful to decide which taxa should be selected for bioaugmentation,
380 with rhizosphere-, detritusphere- and aging root-specific taxa as candidates. Ideally, organisms
381 from different guilds could be combined into consortia that should effectively contribute more
382 than a single taxon to the nitrogen economy of the plant through mineralization of nitrogen
383 coinciding with plant demand over time.
384

385 **References**

386 1. Moreau D, Bardgett RD, Finlay RD, Jones DL, Philippot L. A plant perspective on nitrogen cycling
387 in the rhizosphere. *Funct Ecol* 2019; **33**: 540–552.
388 2. Myrold, DD. Transformations of nitrogen. *Principles and applications of soil microbiology* 2021;

Elsevier 385-421.

3. Vitousek P, Howarth R. Nitrogen limitation on land and in the sea: How can it occur? *Biogeochemistry* . 1991; **13**: 87-115

4. Harte J, Kinzig AP. Mutualism and competition between plants and decomposers: implications for nutrient allocation in ecosystems. *Am Nat* 1993; **141**: 829–846.

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

6. 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(5)**: 841–853

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

8. 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**: 183–192.

9. 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**: 119–128.

10. 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**: 470–480.

11. Chapin FS, Stuart Chapin F, Matson PA, Mooney HA. Principles of Terrestrial Ecosystem Ecology. 2002. Springer

12. Schimel JP, Bennett J. NITROGEN MINERALIZATION: CHALLENGES OF A CHANGING PARADIGM. *Ecology* . 2004. **85**: 591–602

13. Hill PW, Jones DL. Plant-microbe competition: does injection of isotopes of C and N into the rhizosphere effectively characterise plant use of soil N? *New Phytol* 2019; **221**: 796–806.

14. Roberts P, Stockdale R, Khalid M, Iqbal Z, Jones DL. Carbon-to-nitrogen ratio is a poor predictor of low molecular weight organic nitrogen mineralization in soil. *Soil Biology and Biochemistry* . 2009. **41**: 1750–1752

15. Jilling A, Keiluweit M, Contosta AR, Frey S, Schimel J, Schnecker J, et al. Minerals in the rhizosphere: overlooked mediators of soil nitrogen availability to plants and microbes. *Biogeochemistry* 2018; **139**: 103–122.
16. 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.
17. Nguyen TTH, Myrold DD, Mueller RS. Distributions of Extracellular Peptidases Across Prokaryotic Genomes Reflect Phylogeny and Habitat. *Front Microbiol* 2019; **10**: 413.
18. Pett-Ridge J, Shi S, Estera-Molina K, Nuccio E, Yuan M, Rijkers R, et al. Rhizosphere Carbon Turnover from Cradle to Grave: The Role of Microbe–Plant Interactions. *Rhizosphere Biology: Interactions Between Microbes and Plants* . 2021. Springer, Singapore, 51–73
19. Zhu B, Cheng W. Rhizosphere priming effect increases the temperature sensitivity of soil organic matter decomposition. *Global Change Biology* . 2011. **17(6)**: 2172–2183
20. Keiblinger KM, Schneider T, Roschitzki B, Schmid E, Eberl L, Hämmerle I, et al. Effects of stoichiometry and temperature perturbations on beech leaf litter decomposition, enzyme activities and protein expression. *Biogeosciences* 2012; **9**: 4537–4551.
21. Tian K, Kong X, Yuan L, Lin H, He Z, Yao B, et al. Priming effect of litter mineralization: the role of root exudate depends on its interactions with litter quality and soil condition. *Plant Soil* 2019; **440**: 457–471.
22. Jackson LE, Strauss RB, Firestone MK, Bartolome JW. Plant and soil nitrogen dynamics in California annual grassland. *Plant Soil* 1988; **110**: 9–17.
23. 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.
24. Brzostek ER, Finzi AC. Substrate supply, fine roots, and temperature control proteolytic enzyme activity in temperate forest soils. *Ecology* 2011; **92**: 892–902.
25. Jaeger CH 3rd, Lindow SE, Miller W, Clark E, Firestone MK. Mapping of sugar and amino acid

441 availability in soil around roots with bacterial sensors of sucrose and tryptophan. *Appl Environ*
442 *Microbiol* 1999; **65**: 2685–2690.
443 26. Herman DJ, Johnson KK, Jaeger CH, Schwartz E, Firestone MK. Root Influence on Nitrogen
444 Mineralization and Nitrification in Avena barbata Rhizosphere Soil. *Soil Sci Soc Am J* 2006; **70**:
445 1504–1511.
446 27. DeAngelis KM, Lindow SE, Firestone MK. Bacterial quorum sensing and nitrogen cycling in
447 rhizosphere soil. *FEMS Microbiol Ecol* 2008; **66**: 197–207.
448 28. Nuccio EE, Starr E, Karaoz U, Brodie EL, Zhou J, Tringe SG, et al. Niche differentiation is spatially
449 and temporally regulated in the rhizosphere. *ISME J* 2020; **14(4)**: 999-1014.
450 29. Etesami H, Alikhani HA. Co-inoculation with endophytic and rhizosphere bacteria allows reduced
451 application rates of N-fertilizer for rice plant. *Rhizosphere* 2016; **2**: 5–12.
452 30. Rubin RL, van Groenigen KJ, Hungate BA. Plant growth promoting rhizobacteria are more effective
453 under drought: a meta-analysis. *Plant Soil* 2017; **416**: 309–323.
454 31. Li H, Qiu Y, Yao T, Ma Y, Zhang H, Yang X. Effects of PGPR microbial inoculants on the growth
455 and soil properties of Avena sativa, Medicago sativa, and Cucumis sativus seedlings. *Soil Tillage*
456 *Res* 2020; **199**: 104577.
457 32. Griffiths RI, Whiteley AS, O'Donnell AG, Bailey MJ. Rapid method for coextraction of DNA and
458 RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial
459 community composition. *Appl Environ Microbiol* 2000; **66**: 5488–5491.
460 33. Barnard RL, Osborne CA, Firestone MK. Responses of soil bacterial and fungal communities to
461 extreme desiccation and rewetting. *ISME J* 2013; **7**: 2229–2241.
462 34. Huang Y, Niu B, Gao Y, Fu L, Li W. CD-HIT Suite: a web server for clustering and comparing
463 biological sequences. *Bioinformatics* 2010; **26**: 680–682.
464 35. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene
465 recognition and translation initiation site identification. *BMC Bioinformatics* 2010; **11**: 119.
466 36. Rawlings ND, Barrett AJ, Finn R. Twenty years of the MEROPS database of proteolytic enzymes,

467 their substrates and inhibitors. *Nucleic Acids Res* 2016; **44**: D343–50.

468 37. Bushnell B. BBTools software package. *URL http://sourceforge net/projects/bbmap* 2014.

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

471 39. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Liaw WHA, Lumley T, et al. gplots: Various R
472 programming tools for plotting data. 2015.

473 40. Wickham H. ggplot2: Elegant Graphics for Data Analysis. 2016. Springer.

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

477 42. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with
478 the Genome Taxonomy Database. *Bioinformatics* 2019.

479 43. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. A complete domain-
480 to-species taxonomy for Bacteria and Archaea. *Nat Biotechnol* 2020; **38**: 1079–1086.

481 44. Lee MD. GToTree: a user-friendly workflow for phylogenomics. *Bioinformatics* 2019; **35**: 4162–
482 4164.

483 45. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000;
484 **28**: 27–30.

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

487 47. Oksanen J, Kindt R, Legendre P, O'Hara B, Simpson GL, Solymos P, et al. The vegan package.
488 *Community ecology package [http://r-forge r-project org/projects/vegan/]* 2008.

489 48. Myrold DD. 15 - Transformations of nitrogen. In: Gentry TJ, Fuhrmann JJ, Zuberer DA (eds).
490 *Principles and Applications of Soil Microbiology (Third Edition)*. 2021. Elsevier, pp 385–421.

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

493 50. Bowman SM, Free SJ. The structure and synthesis of the fungal cell wall. *Bioessays* 2006; **28**: 799–
494 808.
495 51. Berg G, Eberl L, Hartmann A. The rhizosphere as a reservoir for opportunistic human pathogenic
496 bacteria. *Environmental Microbiology* . 2005; **7**: 1673–1685
497 52. DeAngelis KM, Brodie EL, DeSantis TZ, Andersen GL, Lindow SE, Firestone MK. Selective
498 progressive response of soil microbial community to wild oat roots. *ISME J* 2009; **3**: 168–178.
499 53. Greenlon A, Chang PL, Damtew ZM, Muleta A, Carrasquilla-Garcia N, Kim D, et al. Global-level
500 population genomics reveals differential effects of geography and phylogeny on horizontal gene
501 transfer in soil bacteria. *Proc Natl Acad Sci U S A* 2019; **116**: 15200–15209.

Competing interests: The authors declare that they have no competing interests

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, JPR and MKF conceived and designed the original rhizosphere-detritusphere study, EEN helped to interpret the data and revise the manuscript. JPR and MKF both 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.