Transcription of macromolecular N degrading genes provides a window into 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

Background: Nitrogen (N) is commonly limiting in soil in part because most N is present in macromolecular organic compounds and not directly available to plants. Using extracellular enzymes, the soil microbial community present near roots (rhizosphere) is largely responsible for transforming organic substrates, which yields nitrogen in forms that are plant-available. Microbial genetic potential (genes) and the cycling of inorganic nitrogen pools have been extensively documented and widely explored in rhizospere soil. But there is scarce information on microbial gene expression for macromolecular nitrogen decomposition and how it relates in space and time to life stages of plant roots. Here we use a suite of time-resolved metatranscriptomes from rhizosphere and bulk soil to follow bacterial and fungal extracellular protease and chitinase expression during rhizosphere aging. We also explore the effect of adding plant root litter as a distinct source of macromolecular carbon and nitrogen.

Results: Expression of extracellular proteases in rhizosphere soil increased with plant growth in the absence of litter, whereas chitinase (*chit1*) was highly upregulated in the detritusphere. Structural groups of proteases were dominated by serineproteases, despite the higher abundance of metalloproteases previously observed in soil and aquatic systems. Extracellular proteases of Betaprotebacteria were more highly expressed in the rhizosphere, whereas those of Deltaroteobacteria and Fungi responded strongly to the presence of litter. We identified distinct functional guilds of taxa specializing in decomposition of proteins in the rhizosphere, detritusphere and in the vicinity of aging roots. We also identify a sub-guild that appears to specialize in protein decomposition in the presence of growing roots as well as litter that

increases its activity in aging rhizosphere soil, which may be particularly N limited. Taxonomic membership of these guilds partially overlapped with functional guilds identified by carbon decomposing enzymes, making up multifunctional guilds.

Conclusions: We found organic N decomposers have distinct spatiotemporal preferences that are selected during plant growth. Rhizosphere microbes which could potentially benefit plants by degrading macromolecular nitrogen were not comparatively active in bulk soil, even in the presence of root litter, implying that they are enhanced near plant roots and less competitive in bulk soils.

Keywords

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

Background

Plants are commonly limited by nitrogen (N) in temperate soils since their access to the largest soil N pools is constrained by the activity of microbes responsible for mineralization of macromolecular nitrogen and N_2 fixation [1–3]. To a large degree, plants depend on microbial degradation of macromolecular forms of N such as proteins and chitin [4–6] and can potentially increase N availability by releasing exudates that stimulate microbial turnover of organic N pools (via 'priming') [7–10]. Depolymerization of high molecular weight detrital organic N is one of the rate-limiting steps in soil N mineralization [11,12] and depends on the activity of extracellular enzymes such as lysozyme, protease, chitinase, nuclease and urease. The resulting N monomers can be further broken down to NH_4 or in some cases taken up directly by the plant [13]. While the genetic potential and bulk-scale activity of extracellular N degrading enzymes is well established, the spatial and temporal dynamics of gene expression underpinning N mineralization are ill-defined. In addition, soil microorganisms capable of degrading macromolecular N are highly diverse [14], and it is unclear when and where specific taxonomic groups are active [15].

Patterns and controls of soil N depolymerization and mineralization may be highly dependent on the habitat (e.g., rhizosphere, detritusphere), since availability of organic-N substrates varies in both time and space, as does the prevalence of fungi, fauna and bacterial degraders, which can enable or limit N-cycling. In the rhizosphere, organic N is available as amino acids, nucleotides, niacin, and choline (derived from plant exudates [10]), lignoproteins and aromatics from sloughed off cells, nucleic acids, and microbial cell wall amino sugar

83 polymers (N-acetylglucosamine, N-acetylmuramic acid) derived from the bloom of cells that
84 develops as roots grow [15]. Moreover, root exudates can drive destabilization of mineral-
85 associated organic matter (MAOM), creating an additional source of macromolecular N [13].
86 Prior work indicates that root exudates can increase degradation of soil organic matter by up to
87 380% [16]. Chitinase and protease enzyme activities were shown to be significantly upregulated
88 in the rhizosphere of wild oatgrass (*Avena spp.*), and N mineralization was highly dependent on
89 quorum sensing [17]. This contrasts with the detritusphere, which has a relatively higher
90 proportion of aromatic N and lignoproteins. Protease enzyme activity has been shown to be
91 enhanced by litter addition [18] and root exudates [19], as well as vary with type of plant litter
92 [20,21]. Obviously, these spatial habitats (rhizosphere and detritosphere) can overlap—as when a
93 root initially grows through bulk soil, or when a growing root begins to age and resemble
94 detritus.

95 Soil N pools and by proxy, N mineralization activity, change with time [22] in part due to
96 the succession of both exudate quality and microbial communities [10,23]. At the bulk soil level,
97 activity of extracellular proteases is thought to drive soil N cycling [12], but what controls this
98 activity is not always clear [24]. Subtle distinctions caused by time and soil habitat are likely
99 masked because most studies are conducted on whole soils, which contain a mixture of
100 rhizosphere, detritusphere and bulk regions. In grassland soils, the rhizosphere is a particularly
101 critical hotspot for microbial biomass and activity due to rapid recycling of root exudation and
102 debris, and functions as a quasi-digestive system decomposing molecules inaccessible to plants.
103 However, root exudation, which serves as a carbon source for rhizosphere bacteria, declines near
104 older root sections [25]. Therefore, it has been hypothesized that bacteria in a mature rhizosphere
105 environment may be forced to target less labile, higher C:N sources of organic N such as plant

106 litter [26]. Indeed, it has been shown that the mature *Avena* sp. rhizosphere has an increased
107 organic N demand relative to the young rhizosphere, where the mature rhizosphere has higher
108 rates of gross nitrogen mineralization and ammonia consumption [28].

109 Of the vast number of microorganisms that exist in a soil, at any given time, only a small
110 proportion (e.g., 8%) are active in the rhizosphere [15,17,27,28]. Gene expression studies show
111 that soil bacteria, in the rhizosphere and in the presence of plant litter, form functional guilds
112 [29–31]. However, there has been little focus on guilds based on macromolecular N degradation.
113 To maintain cellular C:N stoichiometry, microbial carbon acquisition must be accompanied by N
114 uptake. Therefore, we postulate that an analogous functional guild organization occurs in
115 expression of extracellular proteases as with CAZy guilds, as exoproteases are diverse enzymes
116 that target a variety of substrates and have also been shown to be influenced by soil
117 characteristics [32–34]. However, while a general balance of C and N degradation is expected to
118 maintain cellular C:N, it does not follow that the same taxa should be involved in both processes.
119 The capacity to utilize specific C and N forms is taxon-specific; e.g., some bulk soil bacteria may
120 degrade carbohydrates as a C source and take up inorganic N, whereas some rhizosphere bacteria
121 may take up simple root exudates as a C source but degrade proteins as an N source.

122 We hypothesized that community-wide expression and taxon-specific differential
123 expression of extracellular proteases and chitinases differed in rhizosphere vs. bulk soil during
124 root growth and that expression profiles would be altered by root litter amendments. To address
125 that hypothesis, we used metatranscriptomes collected over time from rhizosphere microcosms,
126 using a robust, well-developed experimental approach [17,23,26,29] with a well-characterized
127 annual grassland soil from Northern California. Mature wild oat grass *Avena fatua* [22] plants
128 were monitored for 3 weeks to determine the role of diverse indigenous rhizosphere bacteria in

129 macromolecular nitrogen degradation by extracellular proteases and chitinases throughout root
130 growth and aging. In addition, we compared functional guilds based on extracellular protease
131 expression with guilds defined by expression of carbohydrate active enzymes (CAZy) capable of
132 degrading root litter.
133
134 **Methods**
135
136 *Experimental design, sample collection and sequencing*
137
138 The experimental design, sample collection and sequence data processing are described
139 in detail in [29]. Briefly, common wild oat *Avena fatua* was grown for six weeks in rhizobox
140 microcosms containing soil from the Hopland Research and Extension Center (HREC) in
141 northern California, a Bearwallow–Hellman loam (pH 5.6, 2% total C) packed at field bulk
142 density. Roots were then allowed to grow into a sidecar soil region with a transparent wall,
143 where the root growth timeline was marked at 3 days, 6 days, 12 days and 22 days. In half of the
144 sidecars, the soil was amended with dried *A. fatua* root detritus (C:N = 13.4) chopped to 1 mm.
145 Bulk soil bags, inaccessible to roots, were placed in each sidecar; a bulk soil treatment amended
146 with root detritus was also included. At each timepoint, three replicate microcosms were
147 destructively harvested for paired rhizosphere and bulk soil, both detritus amended and
148 unamended, yielding a total of 48 samples representing four treatments at four timepoints:
149 rhizosphere, rhizosphere + litter, bulk, and bulk + litter.
150 Harvested soil (1 g) was placed immediately in Lifeguard Soil Preservation Reagent
151 (MoBio) and processed according to the company protocol. Roots and supernatant were

152 removed, and the soil was stored in -80°C. DNA and RNA were co-extracted using a phenol-
153 chloroform procedure [35,36] and separated with an AllPrep kit (Qiagen). RNA was DNase
154 treated (TURBO DNase, Thermo-Fisher Scientific), depleted in ribosomal RNA (Ribo-Zero
155 rRNA Removal Kit, Epicentre) and reverse transcribed into cDNA. Metatranscriptomes were
156 sequenced for 48 samples on an Illumina HiSeq 2000 2x150 (TruSeq SBS v3) at the Joint
157 Genome Institute (JGI), in Walnut Creek CA.
158
159 *Expression of chitinase and extracellular protease genes identified in assembled*
160 *metatranscriptomes*
161
162 Raw reads were quality-trimmed (Q20) and rRNA and tRNA reads were removed.
163 Library size, evenness, richness and Shannon diversity were comparable between experimental
164 groups with a mean library size of 43 M paired end reads after filtering. Quality-controlled
165 metatranscriptomic reads were assembled into contigs within each sample [29]. Contigs smaller
166 than 200bp were discarded and the remaining contigs from all samples were clustered at 99%
167 identity with cd-hit-est keeping the longest sequence as the cluster representative [37]. Open
168 reading frames (ORFs) were predicted using Prodigal [38]. Extracellular protease ORFs were
169 identified by reciprocal BLAST to extracellular peptidases from the MEROPS database [14,39].
170 ORFs were assigned a peptidase group (serine-, metallo-, cysteine-peptidase and others) by their
171 best reciprocal BLAST hit. Taxonomy was determined by best BLAST hit to the NCBI non-
172 redundant database (NR, access date July 29th 2019). Chitinase ORFs were identified using six
173 chitinase hidden Markov models (HMMs) from the Kyoto Encyclopedia of Genes and Genomes
174 (KEGG): K01183, K03791, K13381, K17523, K17524 and K17525. Only the first three were

175 detected in our dataset. Reads were then mapped back to ORFs requiring minimum identity 95%
176 and 75% breadth using bbmap [40]. Read counts were normalized using DESeq2 [41]. Heat
177 maps of normalized counts were generated in R using the heatmap2 function in gplots [42].
178 Normalized transcript counts per gene per time point were compared between groups
179 (rhizosphere, litter and litter-amended rhizosphere) using ANOVA and Tukey post-hoc test.
180 Boxplots were generated in ggplot2 [43].
181
182 *Expression of extracellular protease genes from a curated collection of Hopland-soil genomes*
183
184 The quality-controlled reads were mapped using BBsplit [40] requiring 80% identity
185 against a dereplicated reference database of 282 HREC soil genomes including isolates [10],
186 single amplified genomes (SAGs) [29], metagenomic assembled genomes (MAGs) (NCBI
187 PRJNA517182) and stable isotope probing MAGs (SIP-MAGs) [44]. On average, 12.3% (range
188 6.2-31.9%) of the reads per library mapped unambiguously to genomes from the reference
189 database. This additional approach was taken to investigate gene expression within the context of
190 a genome and to search for overlap in guild membership between extracellular protease defined
191 guilds and previously defined CAZy guilds [29]. Three MAGs that were classified as unknown
192 or domain bacteria in previous studies were assigned taxonomy using GTDB-Tk version 1.3.0
193 [45,46]. Verification of the taxonomic assignment of MAG Burkholderiales_62_29 was done
194 using GToTree [47] with complete reference genomes of Betaproteobacteria from RefSeq (Feb
195 28, 2020).
196 Open reading frames (ORFs) were predicted in all genomes using prodigal [38] and
197 annotated using KEGG [48] and ggKbase (http://ggkbase.berkeley.edu). Extracellular proteases,

198 which do not have hidden Markov models (HMMs) capable of separating them from intracellular
199 proteases, were identified by gene nucleotide identity of at least 90% and coverage of at least
200 60% to *de novo* assembled ORFs of extracellular protease from the metatranscriptomes. Gene
201 counts were identified using Rsubread featureCounts [49].
202 Differential expression of chitinases in the genome collection was already performed in a
203 previous publication on expression of carbohydrate active enzymes (CAZy) and was therefore
204 not repeated here [29].
205
206 *Statistical analyses*
207
208 All features and their abundance (represented by metatranscriptomic read counts
209 normalized to sequencing depth) were analyzed with DESeq2 [41] requiring p-value < 0.05
210 (adjusted for multiple comparisons). Ordination and visualization were conducted in R using
211 ggplot2 [43] and vegan [50]; PERMANOVA (vegan function adonis) was used to detect
212 significant treatment factors affecting expression of nitrogen depolymerization genes. We define
213 'guilds' as groups of taxa with similar gene upregulation patterns of extracellular proteases in
214 both time and soil habitat. Guilds were assigned by hierarchical clustering based on differential
215 expression of extracellular proteases compared to unamended bulk soil, generating four
216 functional guilds: "Rhizosphere", "Detritusphere", "Aging root", and "Low response". Effects of
217 the four experimental treatments (bulk, rhizosphere, litter and litter-amended rhizosphere) on
218 protease gene expression were assessed by ANOVA (p adjusted for multiple comparisons), N=3
219 per timepoint.
220

221 **Results**

222

223 *Expression of enzymes targeting macromolecular N*

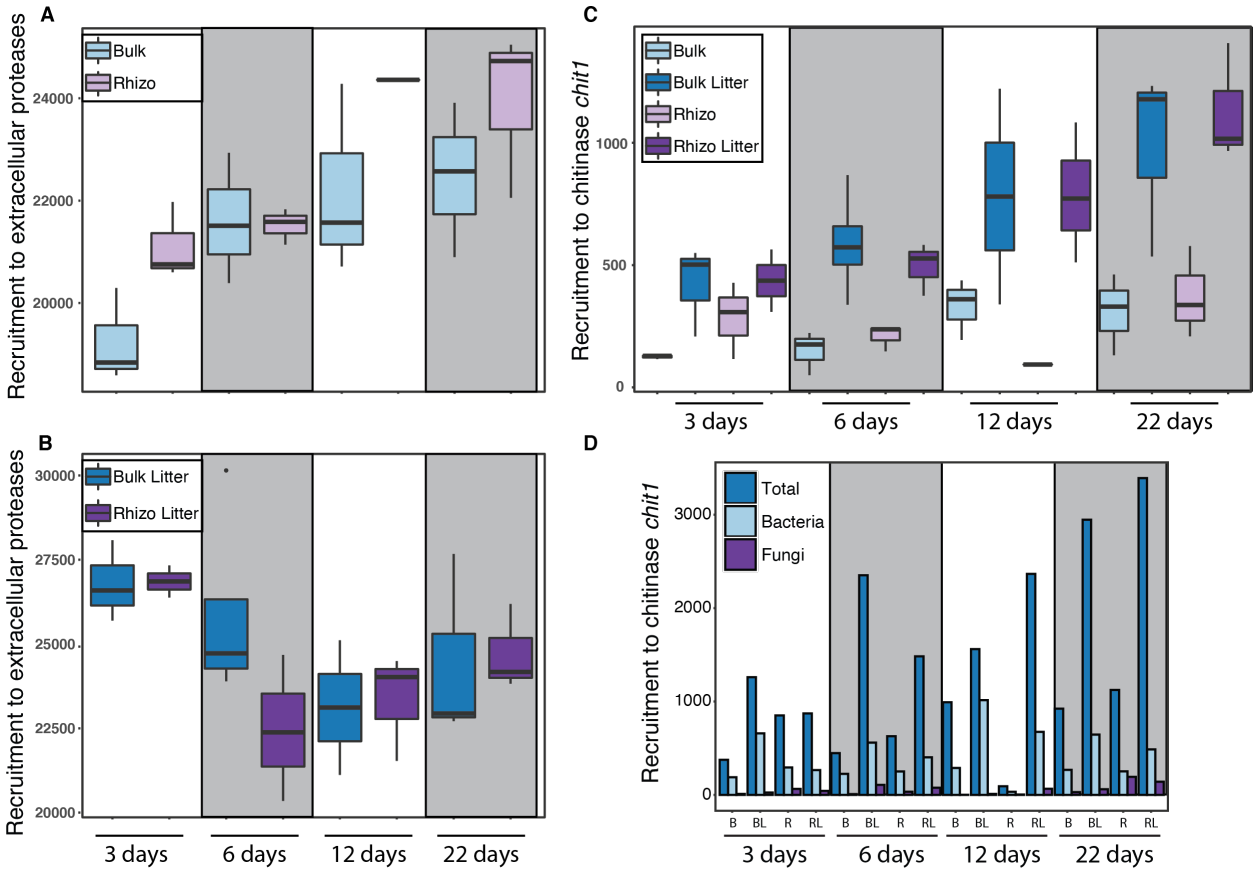
224

225 Multiple genes related to degradation of macromolecular N were identified in this
226 experiment: extracellular nuclease (Xds), urease (UreABC), lysozyme, chitinase (chit1) and
227 extracellular proteases. Expression of extracellular proteases was an order of magnitude higher
228 than the extracellular N degrading enzymes (sup. fig. S1), and like chitinases, was affected by
229 the presence of litter and roots. In contrast, extracellular nuclease (Xds), urease (UreABC) and
230 lysozyme were not influenced by either living or decaying roots, thus, we chose to focus on
231 extracellular proteases and chitinases.

232 Normalized transcript abundance for 4948 extracellular protease genes from bacteria
233 (4846) and fungi (102) were quantified in the sampled soils. Overall, extracellular protease
234 expression was significantly affected by time (3-way PERMANOVA, $F=2.8$, $p=0.038$), litter
235 amendment ($F=118.9$, $p<0.001$) as well as interactions between time:treatment ($F=26.9$,
236 $p<0.001$), time:location ($F=7.7$, $p<0.001$) and location:treatment ($F=15.5$, $p<0.001$). Expression
237 in unamended soil (no litter) generally increased over time (fig. 1A), whereas litter-amended soil
238 had initially high expression that then decreased and leveled off over time (fig. 1B).

239 The 73 distinct chitinase transcripts were expressed at a substantially lower level than the
240 extracellular proteases. The expression of the *chit1* gene increased over time in litter-amended
241 soils and was higher than in the non-litter treatments (3-way PERMANOVA, $p<0.001$) (fig. 1C).
242 Expression of fungal chitinases was lower than bacterial chitinases, and both were generally
243 lower than eukaryotic chitinases (fig. 1D). Transcripts for chitinase KEGG orthologs *CHI3L1_2*,
244 *CHI3L1_4* and *chiA* were not detected at all and the putative chitinase K03791 had low

245 expression and no significant effects with time or treatment (data not shown). Chitinase CHID1
246 was detected but was most closely related to plants and therefore disregarded. A companion
247 study analyzed differential expression of chitinases within our curated genome collection and
248 while only one MAG was found to upregulate chitinase, it did so in the presence of litter [29].
249



250
251

252 Figure 1: Aggregated expression of extracellular proteases in bulk and rhizosphere soils from
253 common wild oat *Avena fatua* microcosms, grown without root litter amendment (A), or with
254 litter amendment (B); expression of chitinase gene *chit1* (C) and mean counts of chitinases at the
255 domain level (D) for unamended bulk soil (B), litter-amended bulk soil (BL), unamended
256 rhizosphere (R) and litter-amended rhizosphere (RL). Boxplots represent 75% of the data,

257 whiskers denote 90% of the data and dots (e.g. panel A bulk litter 6 days) represent outliers. Note
258 that the scale of the Y axis varies between panels.
259
260 *Structural groups of extracellular proteases*
261
262 Structural groups of extracellular proteases can be soil-specific and pH dependent [32].
263 The main groups of extracellular proteases found in our assembled metatranscriptomes were
264 serine-, metallo- and cysteine-proteases (2679, 1949 and 496, respectively out of 5295 variants
265 clustered at 99% identity). Aggregated expression patterns of all variants in each group reflected
266 the same order (serine>metallo>cysteine). Expression of serine- was consistently higher than
267 metallo-protease across all treatments (ANOVA F=2392, p=0) (fig. 2, sup. table S4). Expression
268 of serine-proteases was also significantly greater in the presence of litter compared to no litter
269 (diff=2337, p<0.001), but root litter amendment did not affect expression of metallo- or cysteine-
270 proteases, and there was no significant effect of time or location on these structural groups (sup.
271 table S4).
272

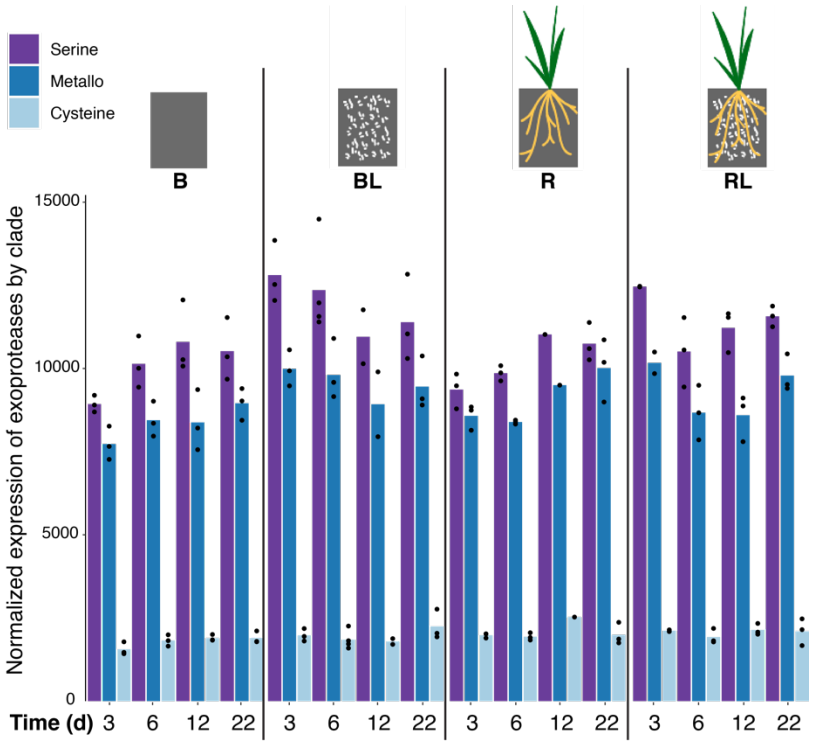


Figure 2: Aggregated normalized expression of extracellular protease structural groups over time in bulk, rhizosphere and root litter-amended soils. Normalized expression per replicate is plotted as dots on top of the bars. Experimental groups shown per time point are (left to right): unamended bulk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

Taxonomy of extracellular proteases

De novo assembled ORFs of extracellular proteases were taxonomically assigned by BLASTP best hit against the NCBI non-redundant database (NR). Since extracellular proteases are not marker genes, we considered taxonomic assignments only at the order level or higher,

287 with the exception of *Rhizobacter*, for which amino acid percent identity values were extremely
288 high (84±8.4% amino acid identity).

289 In the rhizosphere, Betaproteobacteria exoproteases were significantly upregulated
290 (ANOVA, $p<0.05$) (fig. 3A). Of 547 variants of exoproteases from this class, 442 were assigned
291 as Burkholderiales. In contrast, at most timepoints, proteases of Cyanobacteria and
292 Thaumarchaeota were significantly downregulated in the rhizosphere compared to bulk soil (fig.
293 3B; sup. table S1). In multiple other taxonomic groups, proteases were significantly upregulated
294 only in the presence of litter: Fungi (fig. 3C), class Deltaproteobacteria (fig. 3D), as well as
295 highly represented phyla Bacterioidetes and Verrucomicrobia (sup. fig. S2; sup. table S1) and
296 classes Chitinophagia and Gammaproteobacteria (sup. table S2). Interestingly, exoprotease
297 expression declined with time for the predatory bacteria Myxococcales (fig. 3E),
298 Bdellovibrionales and Cytophagia, while clades such as Chloroflexi (fig. 3F) and Actinobacteria
299 had increased exoprotease expression at the final sampling point. While the Actinobacteria and
300 Acidobacteria had a high number of variants and high expression of exoproteases, we did not
301 detect a significant effect of either time or soil habitat/ amendment. Normalized protease
302 expression data are summarized by phylum (sup. fig. S2), class (sup. fig. S3) and order (sup. fig.
303 S4), and ANOVA F and p values are in sup. table S1, sup. table S2 and sup. table S3
304 respectively.

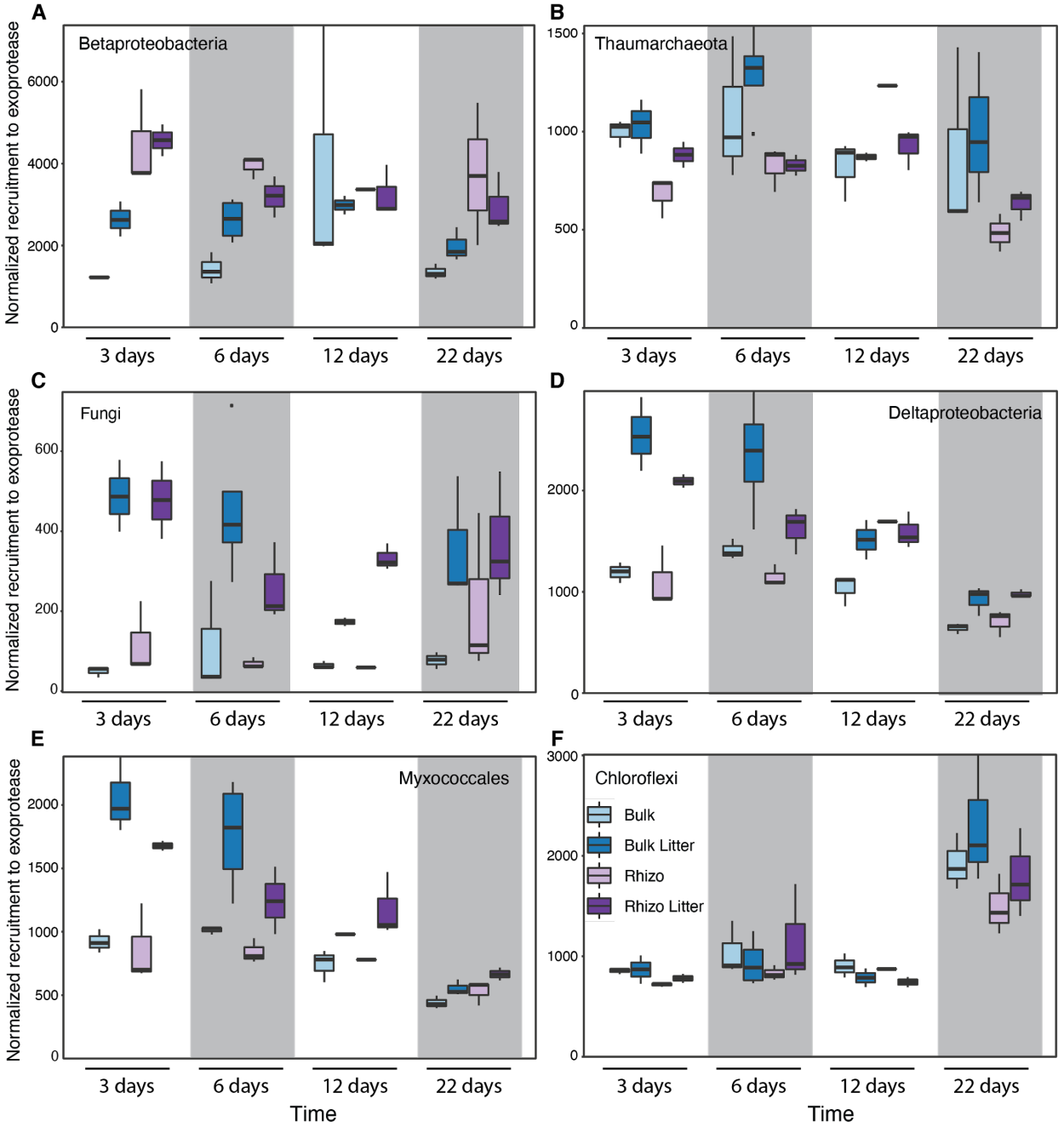


Figure 3: Expression of extracellular proteases over time for select taxonomic groups: (A) Betaprotebacteria (n=547) were upregulated in the rhizosphere, (B) Thaumarchaeota (n=112) were downregulated in the rhizosphere, (C) Fungi (n=99) were upregulated in the presence of litter, (D) Detaproteobacteria (n=228) were upregulated in the presence of litter, (E)

311 Myxococcales (n=147) were upregulated in the presence of litter and decreased over time and (F)
312 Chloroflexi (n=153) increased in the last time point. The legend in panel F applies to all panels.
313 Note that as expression levels varied over taxonomic groups and to emphasize patterns, the y
314 axes do not have the same scale.
315
316 *Functional guilds*
317 To define functional guilds with a population-centric analysis, we mapped transcriptome
318 reads to a collection of 282 genomes and metagenome-assembled genomes generated from the
319 same soil and site [29]. Read counts were used to determine differential expression compared to
320 unamended bulk soil. Each genome contained multiple genes coding for extracellular proteases.
321 Hierarchical clustering of the mean differential expression of extracellular protease genes within
322 each genome was used to define four functional guilds: “rhizosphere”, “detritusphere”, “aging
323 root” and “low response” (fig. 4) following definitions first described in [29]. Similarly,
324 hierarchical clustering based on differential expression of the most highly upregulated *de novo*
325 assembled ORFs compared to unamended bulk soil revealed clear rhizosphere and detritusphere
326 guilds (sup. fig. S5). There was some overlap between guilds identified here and guilds identified
327 based on expression of CAZy genes, but the degree of overlap varied by guild (sup. fig. S6).
328 Within the aging root guild, we noticed one member, representing an aggregated
329 population transcriptome, that increased differential expression of extracellular proteases at the
330 last time point in the presence of both roots and litter more than would have been predicted by
331 the sum of the treatments alone. As this pattern differed from the rest of the guild, we identified
332 this member as a sub-guild labeled “late synergist” due to its expression pattern. This MAG,
333 Burkholderiales_62_29, was identified by two independent phylogenomic analyses as

Rhizobacter (sup. fig. S7). Burkholderiales_62_29 has 16 different extracellular protease genes, all of which had similar significant upregulation patterns. While the sub-guild here contained 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 of the same samples revealed 10 operational taxonomic units (OTUs) of order Burkholderiales which, like *Rhizobacter*, are not assigned a family [29]. Additionally, this sub-guild may include more taxa for which we have no MAGs and therefore could not be included in this analysis.

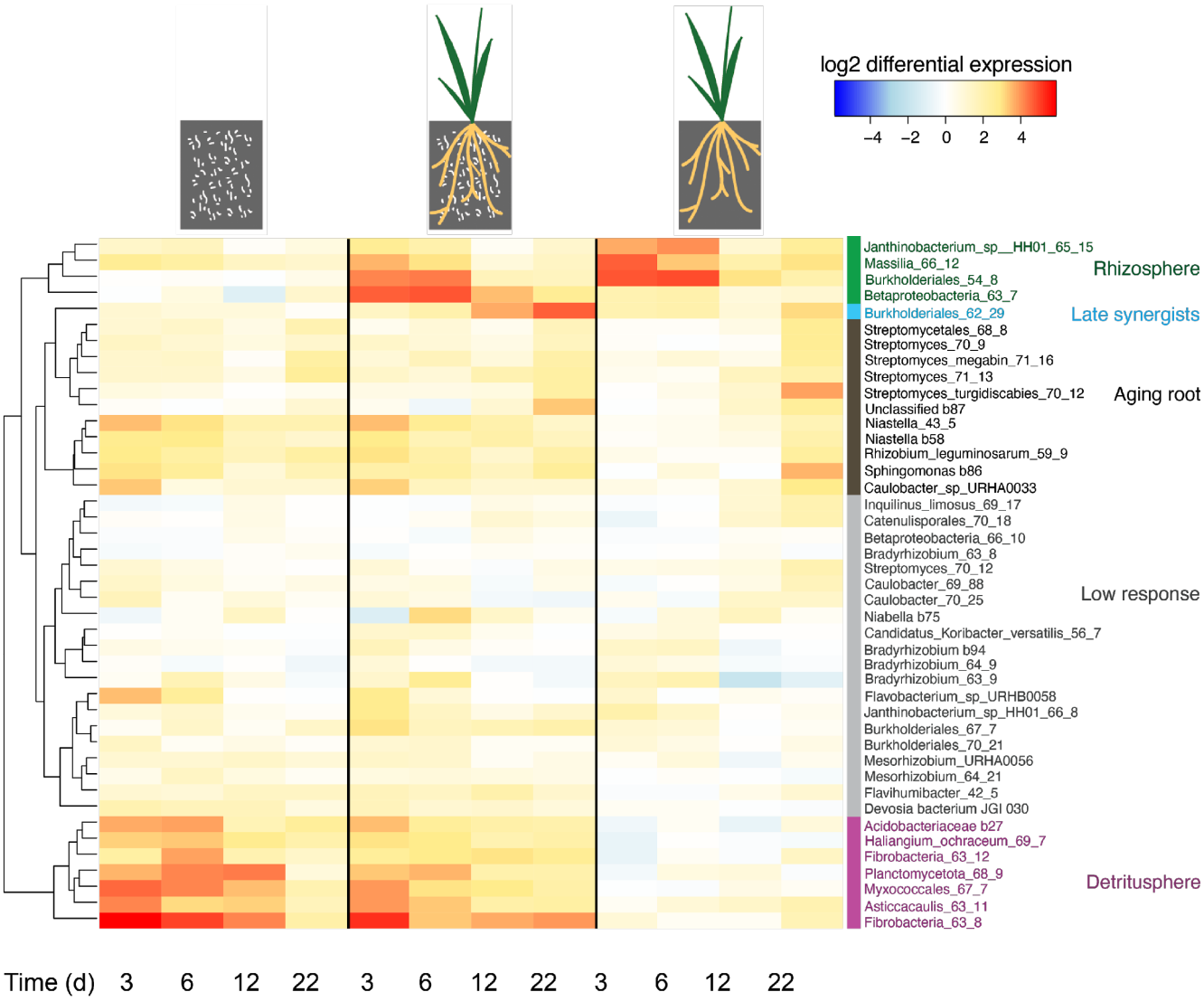


Figure 4: Functional guilds defined by hierarchical clustering of extracellular protease differential gene expression during a 22-day *Avena fatua* microcosm experiment. Each row represents the mean differential expression of extracellular protease genes mapped to a reference genome. Treatments are (left 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 exoprotease 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

Nitrogen in grassland soil is present mostly in the form of organic macromolecules such as proteins and chitin [22,51]. Rhizosphere bacteria can acquire nitrogen through degradation of extracellular macromolecular N, and this process can be enhanced by root exudates especially in unfertilized or N limited soil [1,9,52]. Amino acids resulting from protein degradation are taken up by microorganisms, and roughly 30% of the amino-acid carbon is respired, leading to excretion of excess ammonium which could benefit the plant [53,54]. While various types of extracellular enzymes degrade macromolecules containing N, we focus here on protease and chitinase as they were significantly affected by litter amendment and by proximity to the plant (bulk soil vs. rhizosphere). A previous study based on the same experiment showed that there was very little change in the bacterial or fungal community composition throughout the experiment [29].

366 Chitinase expression, while two orders of magnitude lower than that of proteases,
367 increased over time in litter-amended samples. This effect could imply that the litter attracted
368 saprotrophic fungi and possibly arthropods, the cell walls of which contain chitin [55]. As most
369 of the chitinases were most closely related to eukaryotic enzymes, it is likely that the chitin-
370 containing hyphae were preyed upon by soil fauna. A previous study with *Avena fatua* in the
371 same soil without a preceding addition of chitin showed that chitinase specific activity was
372 highest near root hairs (several days old rhizosphere), medium in bulk soil and lowest near
373 mature roots (one week old rhizosphere) [17]. In contrast, in rhizosphere with no litter addition,
374 we saw little change in chitinase expression over time. The slight differences in these results may
375 simply reflect the fact that the preceding study assayed enzyme activity and the current study
376 assessed gene expression.
377 We found that in unamended soil, expression of extracellular proteases was higher in the
378 rhizosphere compared to bulk soil, potentially resulting from input of carbon from root exudates
379 creating a higher demand for nitrogen as well as microbial competition over inorganic nitrogen
380 with the plant. Indeed, DeAngelis et al. showed that protease specific activity was significantly
381 higher in young rhizosphere compared to bulk soil, but activity was not significantly different
382 based on root age [17]. In litter-amended soil, expression of extracellular protease was highest
383 regardless of location at 3 days, suggesting that litter-added carbon overwhelms the effect of root
384 exudates at this early point in time. At 12 and 22 days there is no difference between litter-
385 amended and unamended soil in either rhizosphere or bulk soil, implying that the effect of litter
386 has waned. Changes in the quantity and quality of rhizodeposits can lead to changes in gene
387 expression over time in the rhizosphere. An increased demand for N is consistent with the higher
388 rates of gross nitrogen mineralization and ammonium consumption found in the rhizosphere of

389 mature compared to young *Avena sp*. [26]. The trend of increase in expression of extracellular
390 protease over time in our unamended rhizosphere soils suggests that as the rhizosphere ages it
391 becomes more nitrogen limited; thus, organic nitrogen would need to be mineralized to support
392 the microbial community.

393 Like many enzymatic reactions, macromolecular N depolymerization can be regulated by
394 substrate availability and end-product concentration [56]. In our study, litter amendment initially
395 led to high expression of extracellular proteases, suggesting that this expression was substrate
396 dependent [57,58] and linked carbon and nitrogen cycling [59]. However, protease activity in
397 grassland soil has also been shown to be affected by a multitude of parameters, such as end-
398 product concentration, aboveground net primary productivity [34] and soil aggregation [33]. We
399 observed an increase over time in expression in unamended samples although no nitrogen or
400 carbon sources were added. We hypothesize that this increase over time may be related to
401 depletion of end products (inorganic nitrogen), leading soil bacteria to invest more in expressing
402 depolymerization enzymes in order to access macromolecular nitrogen.

403 Spatiotemporal patterns of differential expression of genes were used to identify
404 functional guilds in which extracellular proteases were upregulated at a specific location or time.
405 These microbial guilds, which differed in taxonomic composition, underwent distinct
406 successional patterns that appear to be driven by niche conditions in the rhizosphere,
407 detritusphere, aging root and low response. Nuccio et al. [29], analyzing the same experimental
408 dataset, showed that microbial community composition (by 16S-rRNA) in each experimental
409 group changed very little compared to gene expression over the course of this experiment [29].
410 Therefore, the changing expression patterns over time that we observed are likely related to

411 changes in environmental conditions or cues, such as macromolecular N availability and
412 inorganic N availability, as opposed to wholesale shifts in community composition.
413 Carbon and nitrogen cycling by soil bacteria are linked through substrates that contain
414 both as well as by the need to maintain cellular C:N. The upregulation of carbon
415 depolymerization may promote increased access to nitrogen through expression of extracellular
416 proteases under the same conditions, leading to an overlap in guild membership. Such a
417 causative correlation has previously been observed in loam agricultural soil using protease
418 enzyme assays [59]. Therefore, we compared guild membership between guilds previously
419 defined by CAZy expression [29] and guilds defined here by hierarchical clustering of
420 extracellular protease expression. We found that rhizosphere guild membership was similar
421 between CAZy and protease guilds, while the detritusphere guilds had little overlap. All
422 members of the rhizosphere guild defined by extracellular proteases were also assigned to this
423 guild by CAZy expression, and all were Betaproteobacteria known to be associated with
424 rhizosphere soil [23,60,61]. In the sub-guild of "aging root" labeled "late synergist",
425 Burkholderiales bin 62_29, was assigned to the rhizosphere guild by CAZy expression. The
426 diverse detritusphere guild defined by extracellular protease expression included facultative
427 predatory bacteria of the order Myxococcales (Deltaproteobacteria). Predatory bacteria produce
428 proteases to bind to and digest prey bacteria [62] which may be attracted to litter. The small
429 amount of overlap between the protease-defined and CAZy-defined detritusphere guilds could
430 potentially be due to protein/amino acid supply of carbon backbones which reduces the need for
431 degradation of carbohydrates by CAZy genes in extracellular protein degraders. Guild overlap or
432 lack thereof contributes to further defining these multifunctional guilds in enzyme space.

433 Extracellular proteases have several structural types that differ in their active site. For
434 example, metalloproteases require zinc in their active site, therefore their activity may by limited
435 by zinc availability. There is currently very little information on when and where these different
436 types are active in soil, but it has been suggested that their activity might depend on soil
437 characteristics [32]. A recent publication by Nguyen et al. [14] showed that most of the
438 extracellular proteases in soil and aquatic environments are Metallo-, Serine- and
439 Cysteinepeptidases (MP, SP and CP respectively) in decreasing order of abundance. We found
440 these same clusters were dominant in our data, but expression patterns from our study revealed
441 dominance of SP over MP. Many bacterial phyla have dozens of copies of genes for secreted
442 peptidase and these genes are enriched especially in Actinobacteria and Proteobacteria, and more
443 so in soil/aquatic environments as opposed to animal associated bacteria. Specifically,
444 Betaproteobacteria are enriched in SP (mean 13 SP copies out of a mean of 25 extracellular
445 proteases per genome), with published genomes of isolates of Burkholderiales genera such as
446 *Burkholderia*, *Janthinobacterium* and *Massilia* containing more than 50 and up to 155 copies per
447 genome [14]. As Betaproteobacteria are enriched in the *Avena* rhizosphere [23], we might have
448 expected higher expression of SP in this environment compared to bulk soil. Surprisingly, we did
449 not observe that pattern; this could be explained by the high expression of genes from
450 Actinobacteria in all groups, as they share the same enrichment pattern with Betaproteobacteria
451 (SP>MP). Nguyen et al. also showed that while bacteria and archaea may carry genes coding for
452 multiple types of these enzymes within the same genome, the enzyme profile of proteolytic
453 activity was soil-specific [14,32]. Given that we see an expression profile of structural protease
454 groups different from that expected from genomes, this may hint that soil bacteria modify
455 expression of these groups in response to soil chemistry.

456

457 **Conclusions**

458

459 We show that extracellular proteases and chitinases are significantly upregulated by the presence
460 of roots and/or plant litter compared to other macromolecular N degrading enzymes. Patterns of
461 extracellular protease expression differ with taxonomy and can be used to define functional
462 guilds. Three guilds, "rhizosphere", "detritusphere" and "aging root" had membership overlap
463 with guilds previously defined by carbohydrate degrading gene expression based on analysis of a
464 CAZy genes dataset from the same experiment. This suggests that some of the members of these
465 guilds co-express carbohydrate- and protein-degradation enzymes. These guilds also exhibited a
466 pattern of temporal succession throughout plant maturation, where expression of protein
467 degrading enzyme genes in the rhizosphere could be triggered by the host plant through control
468 of exudation, or alternatively, may be a response to competition with the plant for labile nitrogen.
469 The taxonomic resolution provided by metatranscriptomes opens a window to the identification
470 of microbes that could enhance N availability to plants. The identity of these microbes has been
471 previously obscured in bulk analyses such as enzyme assays, rate measurements and even qPCR.
472 It is not far-fetched to imagine augmentation of soil or seeds with bacteria that specialize in
473 macromolecular organic N degradation to reduce the use of fertilizers. Ideally, augmentation
474 should use bacteria endemic to the specific plant rhizosphere. We show here that the identity of
475 such bacteria depends on plant life stage and soil characteristics such as organic material load.

476

477 **References**

478 1. Moreau D, Bardgett RD, Finlay RD, Jones DL, Philippot L. A plant perspective on nitrogen
479 cycling in the rhizosphere. Power S, editor. Funct Ecol. 2019;33:540–52.
480 2. Myrold, DD. Transformations of nitrogen. Principles and applications of soil microbiology.
481 Prentice Hall; 1999; 550.
482 3. Vitousek P, Howarth R. Nitrogen limitation on land and in the sea: How can it occur?.
483 Biogeochemistry. 1991.
484 4. Harte J, Kinzig AP. Mutualism and competition between plants and decomposers:
485 implications for nutrient allocation in ecosystems. Am Nat. 1993;141:829–46.
486 5. Paul EA, Clark FE. Soil microbiology and biochemistry. Academic Press. New York, USA.
487 1996.
488 6. Kinzig AP, Harte J. Selection of micro-organisms in a spatially explicit environment and
489 implications for plant access to nitrogen. Journal of Ecology. 1998. p. 841–53.
490 7. Dijkstra FA, Carrillo Y, Pendall E, Morgan JA. Rhizosphere priming: a nutrient perspective.
491 Front Microbiol. 2013;4:216.
492 8. Zhu B, Gutknecht JLM, Herman DJ, Keck DC, Firestone MK, Cheng W. Rhizosphere priming
493 effects on soil carbon and nitrogen mineralization. Soil Biol Biochem. 2014;76:183–92.
494 9. Meier IC, Finzi AC, Phillips RP. Root exudates increase N availability by stimulating
495 microbial turnover of fast-cycling N pools. Soil Biol Biochem. 2017;106:119–28.
496 10. Zhalnina K, Louie KB, Hao Z, Mansoori N, da Rocha UN, Shi S, et al. Dynamic root exudate
497 chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community
498 assembly. Nat Microbiol. 2018;3:470–80.
499 11. Chapin FS, Stuart Chapin F, Matson PA, Mooney HA. Principles of Terrestrial Ecosystem
500 Ecology. 2002.

501 12. Schimel JP, Bennett J. Nitrogen mineralization: challenges of a changing paradigm. Ecology.
502 2004. p. 591–602.
503 13. Jilling A, Keiluweit M, Contosta AR, Frey S, Schimel J, Schnecker J, et al. Minerals in the
504 rhizosphere: overlooked mediators of soil nitrogen availability to plants and microbes.
505 Biogeochemistry. 2018;139:103–22.
506 14. Nguyen TTH, Myrold DD, Mueller RS. Distributions of Extracellular Peptidases Across
507 Prokaryotic Genomes Reflect Phylogeny and Habitat. Front Microbiol. 2019;10:413.
508 15. Pett-Ridge J, Shi S, Estera-Molina K, Nuccio E, Yuan M, Rijkers R, et al. Rhizosphere
509 Carbon Turnover from Cradle to Grave: The Role of Microbe–Plant Interactions. Rhizosphere
510 Biology: Interactions Between Microbes and Plants. 2021. p. 51–73.
511 16. Zhu B, Cheng W. Rhizosphere priming effect increases the temperature sensitivity of soil
512 organic matter decomposition. Global Change Biology. 2011. p. 2172–83.
513 17. DeAngelis KM, Lindow SE, Firestone MK. Bacterial quorum sensing and nitrogen cycling in
514 rhizosphere soil. FEMS Microbiol Ecol. 2008;66:197–207.
515 18. Keiblinger KM, Schneider T, Roschitzki B, Schmid E, Eberl L, Hämmerle I, et al. Effects of
516 stoichiometry and temperature perturbations on beech leaf litter decomposition, enzyme
517 activities and protein expression. Biogeosciences. Copernicus GmbH; 2012;9:4537–51.
518 19. Tian K, Kong X, Yuan L, Lin H, He Z, Yao B, et al. Priming effect of litter mineralization:
519 the role of root exudate depends on its interactions with litter quality and soil condition. Plant
520 Soil. Springer Science and Business Media LLC; 2019;440:457–71.
521 20. Mukhopadhyay S, Joy VC. Influence of leaf litter types on microbial functions and nutrient
522 status of soil: Ecological suitability of forest trees for afforestation in tropical laterite wastelands.
523 Soil Biol Biochem. Elsevier; 2010;42:2306–15.

21. Srilakshmi A, Dakshayani R, Saigopal DVR, Narasimha G. Influence of forest litter on soil enzyme activities. Ecol Environ Conserv. researchgate.net; 2012;18:105–12.

22. Jackson LE, Strauss RB, Firestone MK, Bartolome JW. Plant and soil nitrogen dynamics in California annual grassland. Plant Soil. Springer; 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 availability in soil around roots with bacterial sensors of sucrose and tryptophan. Appl Environ Microbiol. 1999;65:2685–90.

26. 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. Madison, WI: Soil Science Society; 2006;70:1504–11.

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

28. Vieira S, Sikorski J, Dietz S, Herz K, Schrumpf M, Bruelheide H, et al. Drivers of the composition of active rhizosphere bacterial communities in temperate grasslands. ISME J. 2020;14:463–75.

29. 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;14;999–1014.

30. Yergeau É, Quiza L, Tremblay J. Microbial indicators are better predictors of wheat yield and quality than N fertilization. FEMS Microbiology Ecology. 2020;96(2):fiz205.

547 31. Geddes BA, Paramasivan P, Joffrin A, Thompson AL, Christensen K, Jorrin B, et al.
548 Engineering transkingdom signalling in plants to control gene expression in rhizosphere bacteria.
549 Nat Commun. 2019;10:3430.
550 32. Nguyen T, Kleber M, Myrold DD. Contribution of different catalytic types of peptidases to
551 soil proteolytic activity. Soil Biol Biochem. 2019;138:107578.
552 33. Fukumasu J, Shaw LJ. The role of macro-aggregation in regulating enzymatic
553 depolymerization of soil organic nitrogen. Soil Biol Biochem. 2017;115:100–8.
554 34. Greenfield LM, Hill PW, Seaton FM, Paterson E, Baggs EM, Jones DL. Is soluble protein
555 mineralisation and protease activity in soil regulated by supply or demand? Soil Biol Biochem.
556 2020;150:108007.
557 35. Griffiths RI, Whiteley AS, O'Donnell AG, Bailey MJ. Rapid method for coextraction of
558 DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based
559 microbial community composition. Appl Environ Microbiol. 2000;66:5488–91.
560 36. Barnard RL, Osborne CA, Firestone MK. Responses of soil bacterial and fungal communities
561 to extreme desiccation and rewetting. ISME J. 2013;7:2229–41.
562 37. Huang Y, Niu B, Gao Y, Fu L, Li W. CD-HIT Suite: a web server for clustering and
563 comparing biological sequences. Bioinformatics. 2010;26:680–2.
564 38. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic
565 gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119.
566 39. Rawlings ND, Barrett AJ, Finn R. Twenty years of the MEROPS database of proteolytic
567 enzymes, their substrates and inhibitors. Nucleic Acids Res. 2016;44:D343–50.
568 40. Bushnell B. BBTools software package. URL http://sourceforge net/projects/bbmap. 2014.

569 41. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-
570 seq data with DESeq2. Genome Biol. 2014;15:550.
571 42. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Liaw WHA, Lumley T, et al. gplots:
572 Various R programming tools for plotting data. 2016; R package version 3.0;
573 https://cran.microsoft.com/snapshot/2016-03-30/web/packages/gplots/gplots.pdf
574 43. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer; 2009; ;10:978-0.
575 44. Starr EP, Shi S, Blazewicz SJ, Probst AJ, Herman DJ, Firestone MK, et al. Stable isotope
576 informed genome-resolved metagenomics reveals that Saccharibacteria utilize microbially-
577 processed plant-derived carbon. Microbiome. 2018;6:122.
578 45. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify
579 genomes with the Genome Taxonomy Database. Bioinformatics. 2019; 36(6):1925–7.
580 46. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. A complete
581 domain-to-species taxonomy for Bacteria and Archaea. Nat Biotechnol. 2020;38:1079–86.
582 47. Lee MD. GToTree: a user-friendly workflow for phylogenomics. Bioinformatics.
583 2019;35:4162–4.
584 48. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res.
585 2000;28:27–30.
586 49. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for
587 alignment and quantification of RNA sequencing reads. Nucleic Acids Res. 2019;47:e47.
588 50. Oksanen J, Kindt R, Legendre P, O'Hara B, Simpson GL, Solymos P, et al. The vegan
589 package. Community ecology package [http://r-forge r-project org/projects/vegan/]. 2008.
590 51. Olk DC. Organic Forms of Soil Nitrogen. Nitrogen in Agricultural Systems. 2008;57–100.

591 52. Cheng W, Parton WJ, Gonzalez-Meler MA, Phillips R, Asao S, McNickle GG, et al.
592 Synthesis and modeling perspectives of rhizosphere priming. New Phytol. 2014;201:31–44.
593 53. Hill PW, Jones DL. Plant-microbe competition: does injection of isotopes of C and N into the
594 rhizosphere effectively characterise plant use of soil N? New Phytol. 2019;221:796–806.
595 54. Roberts P, Stockdale R, Khalid M, Iqbal Z, Jones DL. Carbon-to-nitrogen ratio is a poor
596 predictor of low molecular weight organic nitrogen mineralization in soil. Soil Biology and
597 Biochemistry. 2009;41(8):1750–2.
598 55. Bowman SM, Free SJ. The structure and synthesis of the fungal cell wall. Bioessays.
599 2006;28:799–808.
600 56. Geisseler D, Horwath WR, Joergensen RG, Ludwig B. Pathways of nitrogen utilization by
601 soil microorganisms – A review. Soil Biology and Biochemistry. 2010. p. 2058–67. Available
602 from: http://dx.doi.org/10.1016/j.soilbio.2010.08.021
603 57. Allison SD, Vitousek PM. Responses of extracellular enzymes to simple and complex
604 nutrient inputs. Soil Biology and Biochemistry. 2005;37(5):937–44.
605 58. Noll L, Zhang S, Zheng Q, Hu Y, Wanek W. Wide-spread limitation of soil organic nitrogen
606 transformations by substrate availability and not by extracellular enzyme content. Soil Biol
607 Biochem. 2019;133:37–49.
608 59. Geisseler D, Horwath WR. Relationship between carbon and nitrogen availability and
609 extracellular enzyme activities in soil. Pedobiologia . 2009;53:87–98.
610 60. Berg G, Eberl L, Hartmann A. The rhizosphere as a reservoir for opportunistic human
611 pathogenic bacteria. Environmental Microbiology. 2005;7(11):1673–85.

612 61. DeAngelis KM, Brodie EL, DeSantis TZ, Andersen GL, Lindow SE, Firestone MK.
613 Selective progressive response of soil microbial community to wild oat roots. ISME J.
614 2009;3:168–78.
615 62. Pasternak Z, Pietrokovski S, Rotem O, Gophna U, Lurie-Weinberger MN, Jurkevitch E. By
616 their genes ye shall know them: genomic signatures of predatory bacteria. ISME J. 2013;7:756–
617 69.
618

619 **Declarations**

620

621 Ethics approval and consent to participate: not applicable
622 Consent for publication: not applicable
623 Competing interests: The authors declare that they have no competing interests
624

625 Availability of data and material: All data used in this publication, including raw reads and
626 genome collection, has been published previously. Metagenomes assembled MAGs can be found
627 in NCBI PRJNA517182, stable isotope probing MAGs in http://ggkbase.berkeley.edu/, single
628 amplified genomes in IMG under study name Mediterranean Grassland Soil Metagenome and
629 single amplified genomes in IMG, see sup. table S2 in Nuccio et al., 2020 for accession numbers
630 [29]. Raw reads can be found in JGI IMG. For JGI ID's (accession numbers) see supplemental
631 table S1 in Nuccio et al. 2020 [29].
632

633 **Acknowledgements**

634

The authors would like to thank Katerina Estera-Molina, Rachel Hestrin, Rachel Neurath and Nameer Baker for their support and helpful feedback on the analyses. We thank Shenjing Shi for contributing to the initial conceptualization of the experiment. Finally, we thank Evan Starr, Ulas Karaoz, Eoin Brodie, Jizhong Zhou, Susannah G. Tringe, Rex R. Malmstrom, Tanja Woyke and Jillian F. Banfield for assisting in the creation of the genome collection.

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

657 author was not personally involved, are appropriately investigated, resolved, and the resolution

658 documented in the literature.