Dual role of humic substances as electron donor and shuttle for dissimilatory iron reduction

Noah Stern[1], Jacqueline Mejia[1], Shaomei He[2,3], Yu Yang[4], Matthew Ginder-Vogel[1], and Eric E Roden*,[2]

[1] Department of Civil and Environmental Engineering, Environmental Chemistry and Technology Program, University of Wisconsin-Madison, Madison, WI 53706, USA

[2] Department of Geoscience, University of Wisconsin-Madison, Madison, WI 53706, USA

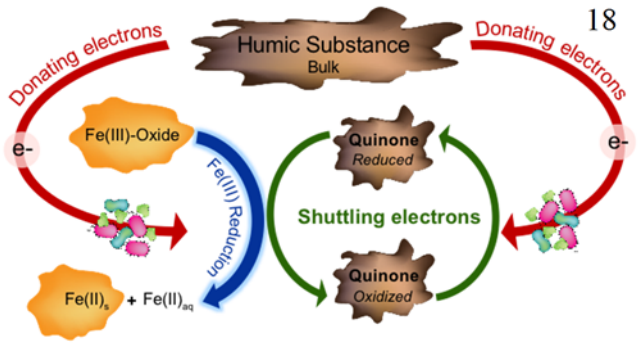
[3] Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53706, USA

[4] Department of Civil and Environmental Engineering, University of Nevada-Reno, Reno, NV, 89557, USA

*Corresponding author. Email: eroden@geology.wisc.edu. Phone: 608-890-7024. Fax: 608-262-0693.

ABSTRACT: Dissimilatory iron-reducing bacteria (DIRB) are known to use humic substances (HS) as electron shuttles for dissimilatory iron reduction (DIR) by transferring electrons to HS-quinone moieties, which in turn rapidly reduce Fe(III) oxides. However, the potential for HS to serve as a source of organic carbon (OC) that can donate electrons for DIR is unknown. We studied whether humic acids (HA) and humins (HM) recovered from peat soil by sodium pyrophosphate extraction could serve as both electron shuttles and electron donors for DIR by freshwater sediment microorganisms. Both HA and HM served as electron shuttles in cultures amended with glucose. However, only HA served as an electron donor for DIR. Metagenomes from HA-containing cultures had an overrepresentation of genes involved in polysaccharide and to a lesser extent aromatic compound degradation, suggesting complex OC metabolism. Genomic searches for the Porin-Cytochrome Complex involved in DIR resulted in matches to *Ignavibacterium/Melioribacter*, DIRB capable of polymeric OC metabolism. These results indicate that such taxa may have played a role in both DIR and decomposition of complex OC. Our results suggest that decomposition of HS coupled to DIR and other anaerobic pathways could play an important role in global subsurface OC metabolism.

TOC ART



INTRODUCTION

Soil organic carbon (SOC) formation involves the biological, chemical and physical transformation of plant and microbial matter into organic products that vary in size at different stages of decomposition and is comprised of a heterogeneous mixture of cellulose, hemicellulose, lignin, lipids, tannins, proteins, sugar monomers and other organic products.[1,2] Since all of these forms of organic carbon are in principle susceptible to microbial degradation, a combination of matrix protection, microbial ecology, enzyme kinetics and geochemical conditions control how SOC is formed and preserved.[3] Materials produced through the degradation of these compounds have been commonly referred to as "humic substances" (HS), organic compounds that have undergone transformation such that they no longer resemble their original physicochemical state.[2,4] "Humic substances" have traditionally been recovered from soils via alkali extraction, which upon further processing leads to operationally defined fractions including humic acid (referred to here as HA), fulvic acid, and humin (referred to here as HM).[5] The term "alkali-extracts" has been recently introduced to explicitly acknowledge that these forms of organic carbon are the result of standardized chemical extraction procedures and not necessarily representative of a native organic materials in soils or sediments.[6]

The significance of SOC in the global carbon cycle is obvious as this reservoir contains more carbon than global vegetation and the atmosphere combined.[7] In addition to serving as a large reservoir of organic carbon (OC), under anoxic conditions microorganisms can respire by transferring electrons to dissolved or solid phase quinone moieties in HS.[8,9] Microbial reduction of quinone moieties accelerates reduction of Fe(III) oxides via a process called electron shuttling, whereby microbially reduced quinones rapidly react to transfer electrons onto Fe(III) oxides.[8-13] This process has received a great deal of attention owing to the broad impact of dissimilatory

iron reduction (DIR) on the geochemistry of anoxic soils and sediments.[14] The same pool of HS that contains quinone moieties also represents a large reservoir of OC that can potentially donate electrons to DIR itself. However, the metabolism of HS under Fe(III)-reducing conditions has not been previously explored. Rather, virtually all previous studies of DIR in the presence of HS have focused on the role of HS as electron shuttles in the presence of excess labile OC.[8,9,15] Coupling of HS degradation to DIR has the potential to create complexity in SOC degradation pathways, since it has recently been shown that DIR can lead to the release of dissolved organic carbon (DOC) associated with Fe(III) oxide surfaces, with important implications for Fe-C associations in soils and sediments.[16,17]

It is well known that hydrolytic and fermentative microorganisms are responsible for breaking down complex OC into organic compounds that can then be utilized by dissimilatory iron reducing bacteria (DIRB).[18] However, the metabolic activity of DIRB such as *Geobacter* is more complex, as such taxa are able to utilize monoaromatic compounds for growth.[19-21] In addition, DIRB from the *Ignavibacteriales* (*Ignavibacterium/Melioribacter*) are capable of degrading complex polysaccharides.[22,23] Thus, the potential exists for DIRB to participate directly in the metabolism of HS, in addition to oxidizing the products of hydrolytic and fermentative metabolism.

Through a combination of geochemical and metagenomic sequencing data this study sought to disentangle electron shuttling and electron donating interactions involving HS and Fe(III)-reducing microbial communities from freshwater sediments. The potential for microbial communities to metabolize HS was determined by quantifying the presence of genes coding for carbohydrate active enzymes (CAZymes) such as glycoside hydrolases, polysaccharide lyases, glycosyl transferases, carbon binding modules and carbohydrate esterases, and genes involved in

the degradation of aromatic carbon.[24] The presence of extracellular electron transfer (EET) systems that are involved in DIR was based on the normalized abundance of porin-cytochrome complex (PCC) and PilA genes first discovered in *Geobacter*.[25-28] Investigating the reduction of Fe(III) oxides and the degradation of HS under controlled conditions provides insight into the microbial community structure and the metabolic pathways involved in coupled Fe-C cycling in soils and sediments.

MATERIALS AND METHODS

Humic Substances. Two different HS obtained from an organic rich histosol from Amherst, Massachusetts were employed in our experiments: a humic acid (HA) fraction representing the combination of two extractions with 0.1M $Na_4P_2O_7$; and the final non-extractable humin (HM) fraction. The chemical properties of the isolated HS used in this study have been previously described.[29,30] The $Na_4P_20_7$ solution used to recover HA differs from the classical pH 13 sodium hydroxide extraction[31], targeting materials associated with organo-metal complexes[32] as opposed to bulk alkali-soluble SOC.

DIR Experiments. DIR experiments were conducted in anoxic 10 mM PIPES (1,4-piperazinediethanesulfonic acid) buffered medium containing 100 μM NH_4Cl and10 μM KH_2PO_4. Duplicate reactors were amended with 100 mmol L^{-1} of reagent-grade hematite (Fe_2O_3, Fisher Scientific) as the Fe(III) oxide mineral phase and 180 mg L^{-1} of either HA or HM as a source of SOC, with or without the presence of 100 μM glucose as an auxiliary electron donor to test the potential for electron shuttling. Note that the HA and HM were added in particulate form to mimic solid-phase SOC. Sterile control reactors containing only 100 μM NH_4Cl, 10 μM KH_2PO_4 and HA/HM showed DOC values below 0.3 ppm after a 72 days. Hematite was chosen as the Fe(III) oxide phase for the sake of consistency with recent studies of biochar-accelerated

Fe(III) oxide reduction.[33] Previous studies have demonstrated its reducibility by freshwater DIRB[34], and unlike ferrihydrite and other Fe(III) oxides, hematite does not undergo reductive phase transformations, which can have a confounding effect on oxide reducibility.[35] The inoculum used for the experiment was obtained from a freshwater pond in Middleton, WI. DIR was determined by monitoring the accumulation of aqueous and 0.5M HCl-extractable Fe(II) using the ferrozine assay.[36]

Metagenome Sequencing and Assembly. Shotgun Illumina sequencing was performed on six metagenomes. On average, 10 Gbp of raw sequences were obtained for each metagenome. Metagenome assembly was performed using the CLC Genomics workbench (version 6.02; CLC bio, Inc., Cambridge, MA, USA). The raw reads were merged to obtain paired-end reads; quality trimmed and filtered by length; and assembled using a *k*-mer of 63 with scaffolding (Table S1). After assembly, the average fold coverage of each contig was estimated and all contigs were uploaded to the Integrated Microbial Genomes with Microbiomes (IMG/MER) database (http://img.jgi.doe.gov/mer) for gene prediction and functional annotation.[37] Genome reconstruction (i.e. binning) from metagenomes was carried out by a coverage-composition algorithm using MaxBin.[38] The completeness and potential contamination of the bins was assessed using CheckM[39] and the resulting high quality bins of interest were further evaluated.

Microbial Community Analysis. Conserved phylogenetic marker and 16S rRNA genes from metagenomes were used to estimate microbial community composition. Phylogenetic marker genes from Phylosift[40] that are present in >99% bacteria and archaea were chosen for this analysis. These genes are represented by 32 COGs (Cluster of Orthologous Groups) listed in Supplementary Table S2. Genes annotated by these 32 COGs were extracted and BLASTP was conducted against the Ref_Seq protein database with an e-value of $1e^{-5}$, with the top 20 hits

being retained. MEGAN4[41] was used for taxonomic classification of the BLASTP results using the lowest common ancestor algorithm.[41] As these phylogenetic marker genes have different lengths due to different gene completeness and different COG families that they represent, the abundance of each phylogenetic marker gene was estimated using its coverage depth weighted by its recovered length and then normalized by its expected full length (i.e. the consensus length of each COG listed in Table S2) according to He et al.[42] The relative abundance of each phylum within the total community was calculated using the normalized abundance of all phylogenetic marker genes.

Identification of OC Degradation Systems. The metagenomes were compared based on functional units of Carbohydrate Active enZymes (CAZymes) or KEGG Orthology (KO). CAZymes are divided into five classes: glycoside hydrolases (GHs); glycosyl transferases (GTs); polysaccharide lyases (PLs); and auxiliary activities (AA) enzymes. The database also includes modules such as the cohesin module, dockerin, the S-layer homology (SLH) module and carbon binding modules (CBMs). CAZyme genes were determined by uploading the annotated metagenomes to a database for automated CAZyme annotation (dbCAN) and using a $1e^{-6}$ cutoff E-value and a 30% cutoff coverage value.[43] When redundancies were detected, classification was determined based on the lowest E-value or highest coverage value. KEGG-KO assignments were performed during IMG/MER gene annotation. We focused on 140 KOs belonging to 22 KEGG modules within the Aromatics Degradation (AD) category.[44-47] Comparison across all six metagenomes was possible by multiplying the scaffold coverage (i.e., scaffold read depth) by a normalization factor, which was computed by dividing 8,420,617,307 bases by the number of bases of raw reads that matched after the assembly of each sample (Table S1). The abundance of genes reported from here on, represents normalized gene copy numbers. To determine the

taxonomic distribution of genes involved in OC degradation, all reads were subjected to BLASTP searches against the NCBI nr database with an E-value cutoff of $1e^{-5}$. BLASTP results were input to MEGAN4 was used for taxonomic classification using the lowest common ancestor algorithm.

Identification of PCC and PilA Genes. PCC gene systems involved in DIRB are composed of one beta-barrel protein and adjacent multi-heme cytochromes[25,26]. These systems were identified by a hidden Markov model (HMM)[48] based on 29 PCC porin gene homologs[26] and created using the hmmbuild function in HMMER 3.0.[49] Confirmation that the identified porin gene contained a beta-barrel protein structure was preformed using PRED-TMBB.[50] Adjacent cytochrome genes and their respective heme-binding sites were searched manually in IMG.[51] The cellular location (periplasmic, outer membrane or extracellular) of all genes was determined by Cello2go.[38] PilA genes[27,28,52] involved in DIRB were identified by IMG protein BLAST using 15 known *Geobacter* specific PilA gene sequences and a maximum E-value of 10^{-5}. The cellular location of the PilA genes was confirmed by Cello2go[38] and gene sequence alignment was performed using PROMALS3D multiple sequence and structure alignment server.[53] Taxonomic assignments of all identified EET gene systems was carried out by BLASTP as described in the previous section.

RESULTS AND DISCUSSION

DIR in the Presence and Absence of HS. The initial rate and long-term extent of hematite reduction were both enhanced by the presence of HA (Figure 1A). In the presence of HA the addition of 0.1 mM glucose stimulated a further increase in DIR reduction; however, the rate and extent of Fe(III) oxide reduction was greater in the presence of HA only compared to glucose only. These results clearly indicate that HA served as a source of electrons for DIRB. In addition, the higher initial rate of DIR in the presence of HA compared to the glucose only reactors is

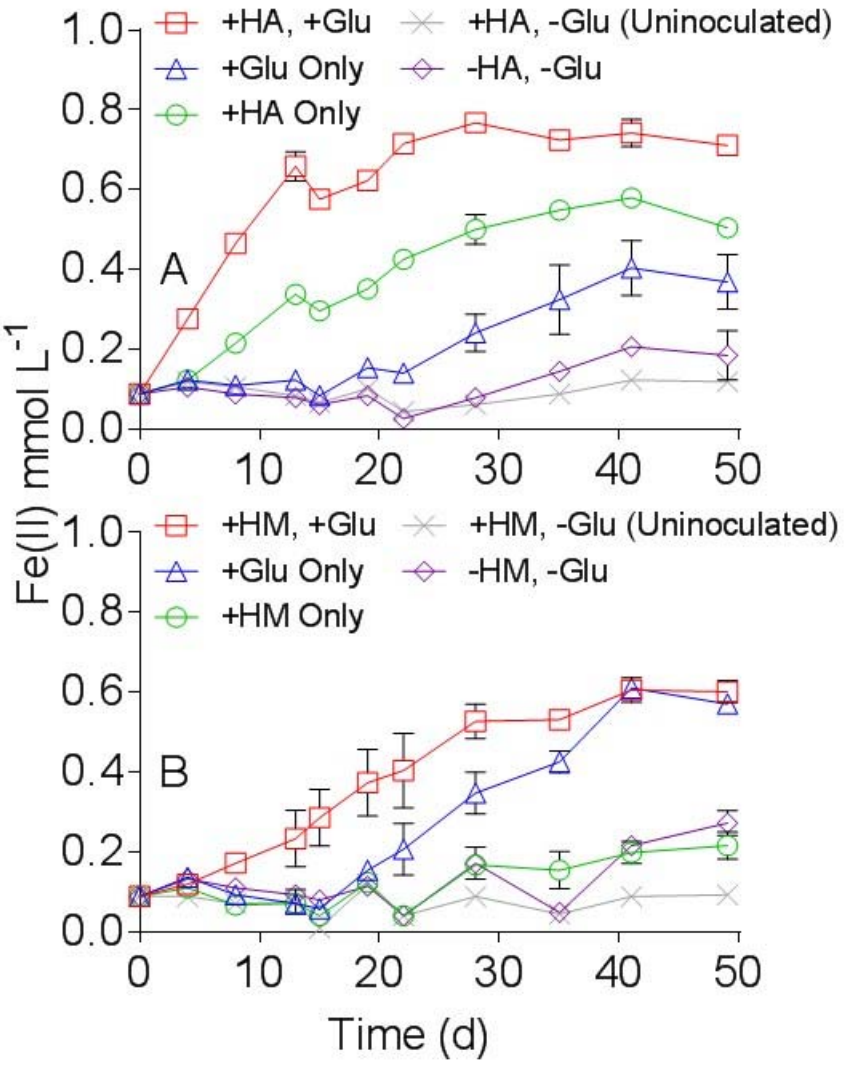


Figure 1. 0.5M HCl-extractable Fe(II) production for the HA (A) and HM (B) hematite reduction experiments. Data points represent the mean ± range of duplicate reactors.

suggestive of an electron shuttling effect analogous to previous studies with solid-phase humic substances.[9] Based on the measured electron-accepting capacity of the HA (2.4 mmol e^- mol C^{-1} [16]) and the amount of OC added with that material (ca. 6.4 mmol C L^{-1}; see other calculations below), the bulk concentration of electron shuttles was on the order of 0.015 mmol e^- L^{-1}. Previous studies with humic acids that have comparable electron-accepting/donating capacity suggest that this amount of electron-shuttling capacity is sufficient to accelerate microbial Fe(III)

oxide reduction.[10]. These calculations also constrain the possible extent of abiotic Fe(III) reduction by residual reducing capacity of the HA (which was not evaluated in this study) to values less than 5% of observed Fe(III) reduction activity in the HA experiments. The total amount of Fe(II) produced in the HA+glucose reactors was equivalent, within experimental error, to the sum of Fe(II) production in the HA-only and glucose-only treatments. The apparent inability of glucose to promote HA metabolism cannot be attributed to an electron acceptor limitation of OC metabolism, because the long-term extent of hematite reduction was 2-fold lower than in previous studies where natural freshwater bacteria were provided with greater amounts of labile OC.[34] This observation suggests that DIR was electron donor limited in our experiments, a conclusion supported by the virtual absence of methane in the headspace of the cultures at the end of the experiment (data not shown).

DIR in the presence of HM alone was similar to that in the no HS addition treatments (Figure 1B, HM only vs. the inoculated control (-HM, -Glu)). As discussed below, the minor amount of reduction in these treatments was likely driven by metabolism of small amounts of OC in the inoculum. The fact that HM in the glucose-amended reactors increased the initial rate of DIR but did not serve as a source of electrons suggests this HM served only as an electron shuttle. Collectively our results show the ability of chemically-extracted humic substances to function as electron shuttles as previously recognized[54], but also in the case of HA as a source of electrons for DIRB.

Microbial Community Composition. Shotgun metagenomic sequencing was carried out to analyze the microbial communities in the cultures and their potential to metabolize complex OC. Conserved phylogenetic marker genes from the metagenomes were used to estimate microbial community composition (Table S3). Reactors containing HA were dominated by a diverse group

of microorganisms including *Spirochaetes*, *Proteobacteria*, *Euryarchaeota*, *Bacteroidetes*, *Ignavibacteriaae*, *Acidobacteria*, *Firmicutes* and *Chloroflexi*. Conversely, *Proteobacteria* and *Acidobacteria* were dominant taxa in both the HM and glucose-only reactors, with the abundance of *Acidobacteria* being higher in the presence of glucose. Overall, HA promoted the development of a more diverse microbial community compared to HM and/or glucose. In particular the high abundance of organisms from phyla such as *Spirochaetes*, *Euryarchaeota*, *Bacteroidetes*, *Ignavibacteriaae*, *Firmicutes* and *Chloroflexi*, all of which are known to degrade complex SOC materials[55-58], implies that these taxa played a role in HA metabolism. Of course, taxa from the *Proteobacteria* and *Acidobacteria* were also substantial components of the HA microbial community and were likely also involved in organic carbon metabolism. In contrast to the HA reactors, HM did not stimulate DIR activity, and did not substantially alter microbial community composition relative to the glucose-only treatments, Based on these results, we infer that HM did not stimulate the growth of any particular phyla thought to participate in the metabolism of HS.

DIR Gene Systems. Organisms responsible for DIR in the cultures were assessed by searching for PCC (Figure 2) and PilA (Figure S2) gene systems. Contrary to expectations, the reactors with highest rates of DIR did not show significant increases in PCC or PilA gene abundance. Experiments with HM and/or glucose had PCC genes mainly assigned to *Geobacter* and *Geothrix*, whereas HA reactors had increases in *Ignavibacterium/Melioribacter* related taxa. The latter taxa are known to be capable of fermentative growth with polysaccharides as well as DIR[22,23,59]; thus the presence of both PCC and CAZyme genes (Figure 3) attributable to these organisms in the HA cultures suggests that they played a role both in DIR and upstream production of fermentation end-products that served as substrates for DIR. It is notable that our

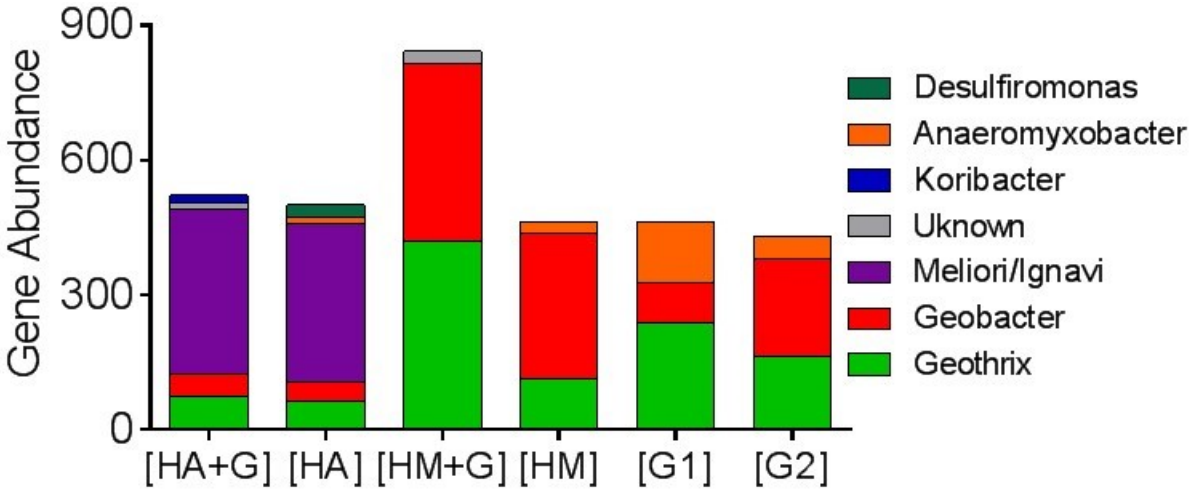


Figure 2. Abundance and taxonomic assignments of porin-cytochrome complex (PCC) genes found in the HA and HM cultures. The gene abundance was normalized by the total assembled base pairs. Abbreviations: HA = Humic Acid; G = Glucose; HM = Humin. G1 and G2 represent replicate glucose-only cultures.

findings reveal for the first time that the previously described DIRB *Geothrix*[60] contains PCC gene systems for reduction of extracellular electron acceptors (see Figure S1 and Supporting Text 1).

General Trends in Organic Carbon Metabolism Gene Abundance. The potential for HA and HM degradation to provide electrons for DIRB was evaluated by determining if genes coding for CAZymes were differentially present in cultures with or without HS at the end of the experiments.[24] A total of 42312 genes across all treatments were classified as one of 258 different CAZyme families. The metagenomes amended with HA had more sequences coding for GHs, GTs, CBMs, CEs, PLs, cohesin, dockerin and SLH compared to the metagenomes from cultures containing HM and/or glucose (Figure 3). Additionally, GHs, GTs and PLs showed slightly higher gene abundance in treatments containing HM compared to those with only glucose (Figure 3). The increase in CAZymes in the HA experiments implies that these enzymes

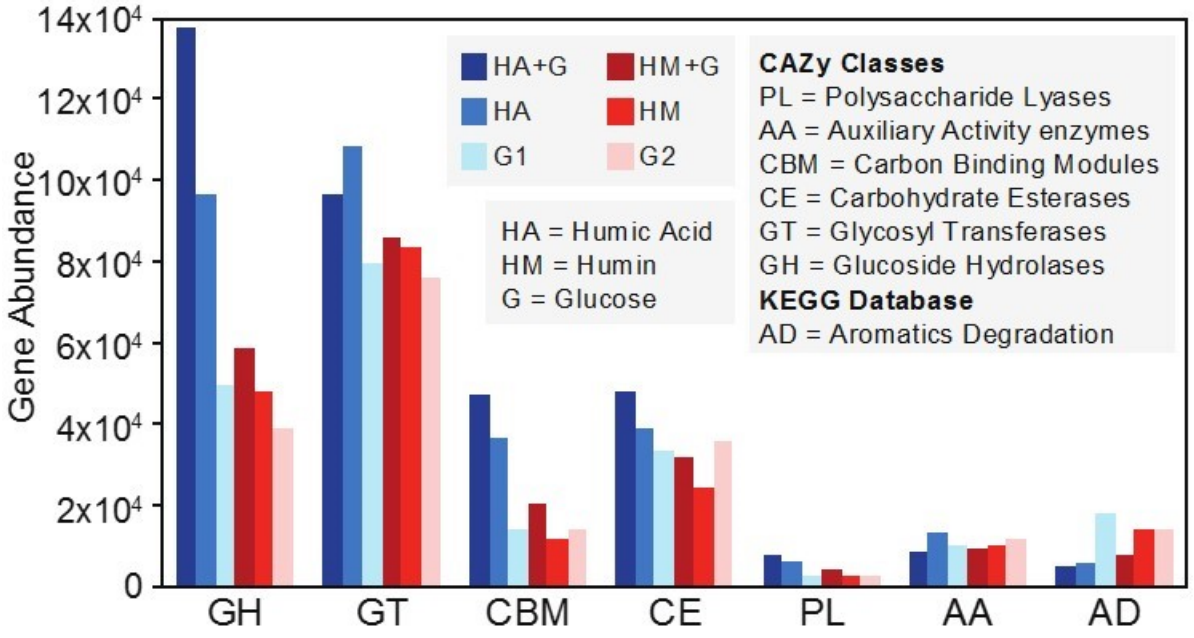


Figure 3. Abundance of genes categorized within a CAZyme class (GH, GT, CBM, CE, PL and AA) and KEGG Aromatics Degradation (AD) category, in the HA and HM data sets. The gene abundance was normalized by the total assembled base pairs.

played an important role in the degradation of HS. CAZyme family gene abundance was similar in both HM-containing and glucose only cultures. These results indicate that HM did not undergo significant enzymatic transformation, which is consistent with the finding that HM did not serve as an electron donor for DIR (Fig. 1B). The significant abundance of CAZymes in HM cultures can be attributed to metabolism of small amounts of complex OC in the inoculum.

The taxonomic origin of each CAZyme was determined by BLASTP. *Bacteroidetes*, *Firmicutes*, *Euryarchaeota*, *Chlorobi*, *Chloroflexi*, *Spirochaetes*, *Planctomycetes*, and unclassified bacteria dominated the taxonomic classification of CAZymes present in HA experiments (Figure 4A). On the contrary, the CAZymes detected in treatments containing HM and/or glucose were dominated by *Proteobacteria*, *Acidobacteria* and unclassified bacteria. GHs, CBMs, CEs and PLs were also dominated by *Euryarchaeota* and *Bacteroidetes* when HM and

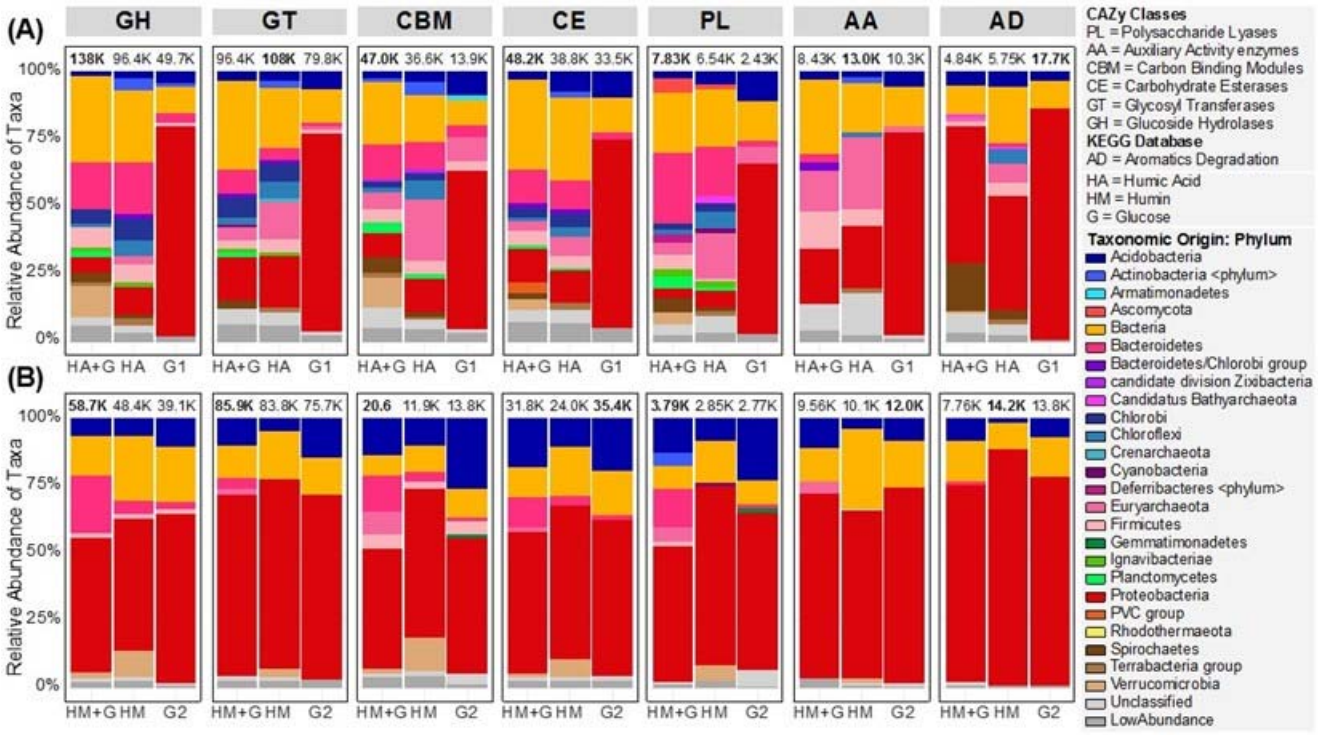


Figure 4. Taxonomic assignment of genes classified within a CAZyme class (GH, GT, CBM, CE, PL and AA) and KEGG Aromatics Degradation (AD) category, in the HA (A) and HM (B) data sets. The y-axis represents the relative abundance of phyla and the number at the top of the bars represents the sum of the gene abundances within a specific category and for a specific experiment. Abbreviations: HA = Humic Acid; G = Glucose; HM = Humin. G1 and G2 represent replicate glucose-only cultures.

glucose were present (Figure 4B). Overall, higher diversity in CAZymes taxonomic assignments present in the HA experiments was expected since the phylogenetic marker genes revealed that these microbial communities were much more diverse (Table S3). It should also be noted that the abundance of unclassified taxa was greater in the HA experiments, suggesting the presence of novel microorganisms capable of complex OC degradation.

The high abundance of GHs, GTs, CBMs, CEs, PLs, cohesin, dockerin and SLH in HA experiments indicates that a large portion of the HA resembles polysaccharides and is in fact

biodegradable. This is in agreement with previous geochemical results and the higher (oxygen + nitrogen)/carbon ratio found in HA (0.94) compared to HM (0.56).[29] Additionally, many of these CAZymes were assigned to *Bacteroidetes*, *Firmicutes*, *Chloroflexi* and *Spirochaetes*, which are important contributors to the degradation of lignocellulosic material in various environments.[61-64] *Planctomycetes* and *Euryarchaeota* were also detected in the presence of HAs and have been associated with cellulose decomposition with GTs[65,66], and metabolism of carbohydrates with GHs[42,63], respectively.

Key Enzymatic Systems Involved in Humic Substance Metabolism. To determine which enzymes had the most important role in the metabolism of HS, we determined which CAZyme families and KEGG modules had gene copy numbers at least two times higher in HS metagenomes compared to glucose only metagenomes. 129 and 72 CAZyme families and, 13 and 6 KEGG KO assignments with this criterion were identified in the HA and HM metagenomes, respectively (see Supporting Text 2 and Tables S4-S7). The CAZyme classes with the highest gene copy numbers were GHs and CBMs, followed by CEs, GTs and PLs (Figure 5). AAs and ADs under this criterion were negligible. The most abundant GHs in HA treatments were categorized as cellulases and endohemicellulases and debranching and oligosaccharide degrading enzymes.[33,42,57,62,67,68] Identical analysis of the HM treatments revealed cellulases, endohemicellulases, and debranching and oligosaccharide degrading enzymes, but also included pectinases and fungal cell wall degrading enzymes. The higher abundance of these CAZyme families in the presence of HA compared to HM and/or glucose treatments is another indication that alkali-extracted SOM contains polysaccharides that are bioavailable for microbial degradation into simpler organic acids utilized by DIRB.[18,20]

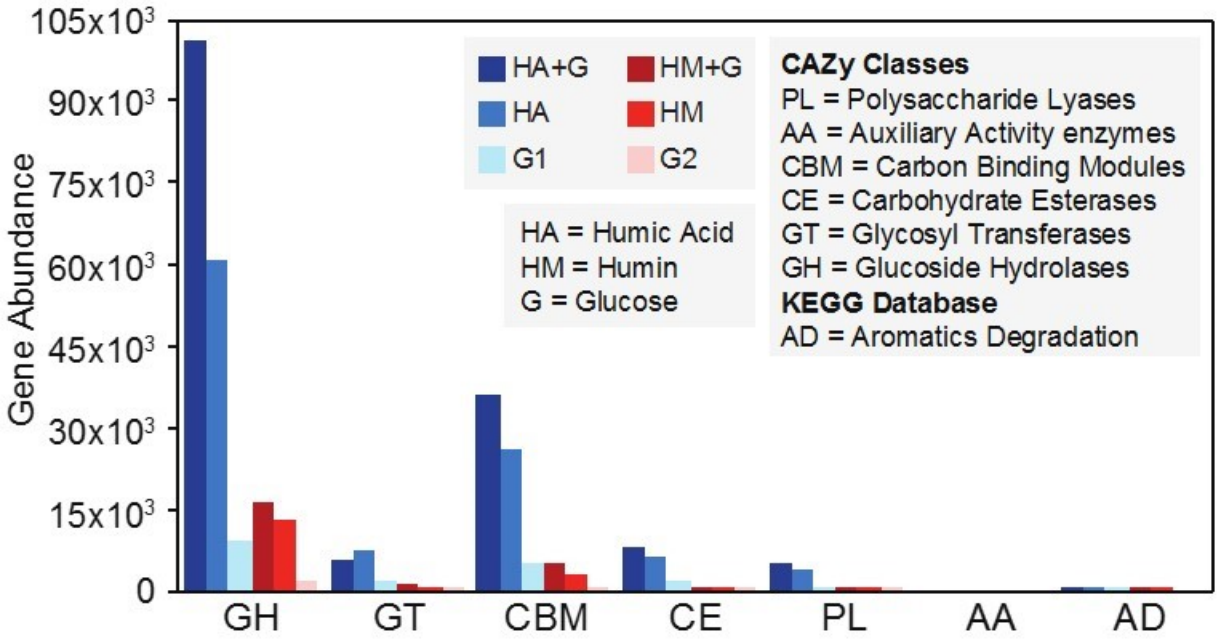


Figure 5. Sum of CAZyme and AD genes that had abundance values at least two times higher in metagenomes amended with HA or HM compared to those with only glucose.

Aromatic Carbon Degradation. Aromatic carbon makes up more than 10% of the bulk carbon in the HA and HM used in this study.[29] This component of HS includes quinone moieties[69] that were likely involved in electron shuttling for DIR, yet these same compounds could undergo degradation coupled to DIR.[3] The abundance of redox enzymes that act in conjunction with CAZymes in aromatic carbon degradation (referred to as AAs) was similar in all metagenomes, but highest when HAs were the only carbon source added (Figure 3). By looking at these trends, one might conclude that AAs were important in the degradation of HAs when glucose is not present. However, AAs were also found in the metagenomes amended with only glucose (Figure 3), suggesting that the inoculum contained aromatic carbon compounds, in addition to other complex OC components as described above, that were subject to degradation. The taxonomic origin of AAs found in the HA treatments (primarily *Firmicutes*, *Proteobacteria* and *Euryarchaeota*) differed from those in the glucose-only and/or HM cultures, which were

dominated by *Proteobacteria* (Figure 4). These results demonstrate that diverse taxa with the capability of aromatic carbon degradation were enriched in the HA cultures.

Since AAs were not predominant in the CAZyme gene inventories, a search for genes involved in the degradation of smaller aromatic compounds using the KEGG Orthology (KO) database was carried out. A total of 2222 genes across all metagenomes were assigned to the Aromatics Degradation (AD) category[44-47], and their abundance was much lower compared to that of GHs, GTs, CBMs and CEs (Figure 3). Most of the AAs and ADs require oxygen to function and their low abundance in our anoxic incubations is not surprising. Moreover, the preferred OC substrate for microorganisms is usually composed of carbohydrates, not aromatic structures.[70] Therefore, the low abundance of AAs and ADs can be explained if the metabolism of more labile OC (i.e., carbohydrates) happens prior to the metabolism of aromatic moieties.

We have illustrated that HS can be broken down by hydrolytic and fermentative microorganisms. However, DIRB such as *Geobacter* and *Ignavibacter* also have the metabolic capacity to utilize monoaromatic compounds for growth.[19-23] Fifteen draft genomes containing a complete EET system were recovered from the metagenomes, and were identified as *Geobacter*, *Geothrix*, *Holophaga*, *Ignavibacteriales*, *Melioribacter*, *Anaeromyxobacter* and *Prolixibacter*. Thirteen genomes within the same taxonomic categories were downloaded from the IMG database in order to have at least one reference for all taxa. All genomes were screened (see Supporting Text 3) for ADs, and while it was confirmed that the *Geobacter metallireducens* reference genome has a complete pathway for toluene degradation and benzoyl-CoA degradation, no such AD degradation pathways were observed in any other genome (Tables S8-S12). Overall there was no correlation between DIR and AD gene abundance (Tables S6 and S7); thus, it seems unlikely that aromatic carbon degradation was a major pathway for HS

296 metabolism in this experiment. However, the large increase in *Ignavibacterium/Melioribacter*
297 related PCC gene systems in the HA vs. HM cultures (Figure 2) does suggest that these
298 organisms played a role in upstream decomposition of complex OC.
299 **Environmental Implications.** Previous studies of the influence of HS on DIR in soils and
300 sediments have focused on electron shuttling effects rather than coupling between DIR and HS
301 (e.g. as a surrogate for SOM) decomposition.[8,9,15] Through a combination of DIR incubation
302 experiments and metagenomic sequence gene counts this study separated the simultaneous
303 processes of electron shuttling and electron donating by HS recovered from peat soil by sodium
304 pyrophosphate extraction. The analysis revealed that metabolism of the HA fraction recovered
305 from the soil was directly coupled to DIR. After a 30 day period roughly 3% of the HA electron
306 equivalents present in the reactors was coupled to DIRB metabolism (see calculation below).
307 Evidence of this phenomenon was provided by the overrepresentation of putative genes coding
308 for enzymes that break down complex lignocellulosic material. For example, 129 CAZy families
309 were at least two times more abundant in experiments containing HA compared to those with
310 glucose only. The potential for microorganisms to utilize a small but significant portion of HA
311 (see below) is consistent with the emerging view of SOC as a continuum of variably
312 decomposable organic compounds.[6]
313 Limited information on the biological degradation of HS under anoxic conditions is
314 currently available,[71] and the potential contribution of HS degradation to overall SOC
315 metabolism under Fe(III)-reducing and other anaerobic respiratory conditions remains an open
316 question. The amount of DIR driven by HA degradation in our experiments was ca. 0.7 mmol
317 Fe(III) L^{-1} in the HA-only treatments (Figure 1A), equivalent to oxidation of ca. 0.175 mmol C
318 L^{-1}. This represents 2.8 % of the total OC added to the reactors, based on the amount of HA

added (180 mg L^{-1}) and the OC content of the HA (42.4 %[72]). Although this may seem like only a minor amount of OC degradation, it is important to consider how much SOC this could represent in a typical OC-rich soil or sediment. Consider a soil with 15% dry weight SOC and a bulk density of 0.4 g dry mass cm^{-3}, typical values for histosols with a significant mineral content[73] such as the one from which the HA used in this study was extracted.[72] The HA fraction of this histosol represented ca. 10% of total SOC[74]. Based on these values, the total HA content of the soil was on the order of 12,000 μmol per cm^3 bulk soil. Assuming, based on our results, that ca. 3% of this material is subject to microbial degradation under anaerobic conditions, this value translates into a biodegradable SOC pool of over 300 μmol cm^{-3}. This represents a surprisingly large pool of labile OC that could drive DIR while at the same time promoting that process via electron shuttling. In situations where Fe(III) oxides are not abundant, this large pool of labile OC would alternatively drive other anaerobic metabolic pathways such as sulfate reduction or methanogenesis. Although, due to the extraction process HS do not directly represent SOM[6], our own and other recent results[71] affirm that chemically-extracted HS represent a pool of SOM that is indeed biodegradable. In addition, recent studies indicate that DIR can liberate OC compounds associated with Fe(III) oxide surfaces[16,17] that are thought to represent a significant pool of preserved SOC in soils and sediments.[75,76] This raises the possibility of a positive feedback loop where DIR driven by SOM depolymerization/fermentation and accelerated by electron shuttling leads to destabilization of Fe-associated OC which in turn further accelerates anaerobic carbon metabolism.

ASSOCIATED CONTENT

Supporting Information. The Supporting Information contains 3 Supporting Text sections, 4 Figures, and 12 Tables that complement the metagenomic results presented in the manuscript,

including a description of the porin-cytochrome complex (PCC) system found in *Geothrix fermentans* (Supplemental Text 1 and Figure S1); results of the PilA gene search (Figure S2); detailed information about the assembly of raw reads using CLC Genomics workbench (Table S1); a list of the phylogenetic marker genes (Table S2) used to determine the microbial community composition (Table S3); an illustration of the presence of carbohydrate active enzyme (CAZyme) modules and their taxonomic assignments (Figures S3 and S4); the abundance and known activity of CAZyme families and KEGG aromatic degradation genes that had an abundance at least two times higher in the presence of HA or HM compared to the glucose only treatments (Supporting Text 2 and Tables S4-S7); and a description and results of screening PCC-containing draft genomes for KEGG aromatic degradation pathways (Supporting Text 3 and Tables S8-S12).

AUTHOR INFORMATION

Corrsponding Author

*Eric E. Roden. Email: eroden@geology.wisc.edu

Phone: 608-890-7024. Fax: 608-262-0693.

Author Contributions

The manuscript was written through contributions of all authors, each of whom has given approval to the final version of the manuscript. N.S. and J.M. contributed equally to writing and editing the manuscript; establishing the conceptual model for dual role of electron donor/shuttle interactions; processing the raw DNA reads (quality trimming and assembly); and annotating the assembled metagenomes. N.S. designed and carried out the experiments; collected the geochemical data; evaluated PCC and PilA gene systems; examined the PCC gene systems and performed preliminary analysis of dbCAN data. J.M. reconstructed genomes (i.e. binning) from

metagenomes and evaluated their completeness and contamination; studied genes involved in the metabolism of complex organic carbon, aromatic compounds and humic acids using the CAZyme and KEGG databases; and determined the presence of aromatic degradation pathways in bins and reference genomes containing PCC gene systems. S.H. used conserved phylogenetic marker and 16S rRNA genes from metagenomes to estimate microbial community composition.

ACKNOWLEDGMENTS

This research was supported by the U.S. Department of Energy, Office of Biological and Environmental Research, Subsurface Biogeochemical Research (SBR) program through grants DE-SC0014275 and DE-SC0016217, and by the SBR Scientific Focus Area at the Pacific Northwest National Laboratory (PNNL). Jacqueline Mejia was supported by the Graduate Engineering Research Scholars and the Biotechnology Training Program (NIH 5T32-GM08349) from the University of Wisconsin-Madison. Funds for the metagenomic sequencing were provided by Office of the Vice Chancellor for Research and Graduate Education at UW-Madison.

REFERENCES

(1) Lorenz, K.; Lal, R.; Preston, C. M.; Nierop, K. G., Strengthening the soil organic carbon pool by increasing contributions from recalcitrant aliphatic bio (macro) molecules. *Geoderma* **2007,** *142*, (1), 1-10.

(2) Schnitzer, M., Humic substances: chemistry and reactions. *Develop. Soil Sci.* **1978,** *8*, 1-64.

(3) Kleber, M., What is recalcitrant soil organic matter? *Environmental Chemistry* **2010,** *7*, (4), 320-332.

(4) MacCarthy, P.; Rice, J.; Aiken, G.; McKnight, D.; Wershaw, R.; MacCarthy, P., *Humic substances in soil, sediment, and water: Geochemistry, isolation and characterization*. Krieger Pub. Co.: 1985; p 708.

(5) Rabenhorst, M. C., and D. Swanson, *Histosols. In M. E. Summer (ed.). Handbook of Soil Science*. CRC press, Boca Raton: 1999.

(6) Lehmann, J.; Kleber, M., The contentious nature of soil organic matter. *Nature* **2015,** *528*, (7580), 60-68.

(7) Intergovernmental Panel on Climate, C., *Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press: Cambridge, 2014.

(8) Lovley, D. R.; Coates, J. D.; Blunt-Harris, E. L.; Phillips, E. J.; Woodward, J. C., Humic substances as electron acceptors for microbial respiration. *Nature* **1996,** *382*, (6590), 445-448.

(9) Roden, E. E.; Kappler, A.; Bauer, I.; Jiang, J.; Paul, A.; Stoesser, R.; Konishi, H.; Xu, H., Extracellular electron transfer through microbial reduction of solid-phase humic substances. *Nat. Geosci.* **2010,** *3*, (6), 417-421.

(10) Jiang, J.; Kappler, A., Kinetics of microbial and chemical reduction of humic substances: implications for electron shuttling. *Environ. Sci. Technol.* **2008,** *42*, (10), 3563-3569.

(11) Kappler, A.; Benz, M.; Schink, B.; Brune, A., Electron shuttling via humic acids in microbial iron (III) reduction in a freshwater sediment. *FEMS Microbiol. Ecol.* **2004,** *47*, (1), 85-92.

(12) Coates, J. D.; Ellis, D. J.; Blunt-Harris, E. L.; Gaw, C. V.; Roden, E. E.; Lovley, D. R., Recovery of humic-reducing bacteria from a diversity of environments. *Appl. Environ. Microbiol.* **1998,** *64*, (4), 1504-1509.

410 (13) Scott, D. T.; McKnight, D. M.; Blunt-Harris, E. L.; Kolesar, S. E.; Lovley, D. R.,
411 Quinone moieties act as electron acceptors in the reduction of humic substances by humics-
412 reducing microorganisms. *Environ. Sci. Technol.* **1998,** *32*, (19), 2984-2989.
413 (14) Lovley, D. R.; Holmes, D. E.; Nevin, K. P., Dissimilatory Fe(III) and Mn(IV) reduction.
414 *Adv. Microb. Physiol.* **2004,** *49*, 219-286.
415 (15) Klüpfel, L.; Piepenbrock, A.; Kappler, A.; Sander, M., Humic substances as fully
416 regenerable electron acceptors in recurrently anoxic environments. *Nat. Geosci.* **2014,** *7*, (3),
417 195-200.
418 (16) Adhikari, D.; Zhao, Q.; Das, K.; Mejia, J.; Huang, R.; Wang, X.; Poulson, S. R.; Tang,
419 Y.; Roden, E. E.; Yang, Y., Dynamics of ferrihydrite-bound organic carbon during microbial Fe
420 reduction. *Geochim. Cosmochim. Acta* **2017,** *212*, 221-233.
421 (17) Zhao, Q.; Adhikari, D.; Huang, R.; Patel, A.; Wang, X.; Tang, Y.; Obrist, D.; Roden, E.
422 E.; Yang, Y., Coupled dynamics of iron and iron-bound organic carbon in forest soils during
423 anaerobic reduction. *Chem. Geol.* **2017,** *464*, 118-126.
424 (18) Lovley, D. R.; Phillips, E. J., Novel mode of microbial energy metabolism: organic
425 carbon oxidation coupled to dissimilatory reduction of iron or manganese. *Appl. Environ.*
426 *Microbiol.* **1988,** *54*, (6), 1472-1480.
427 (19) Lovley, D. R.; Lonergan, D. J., Anaerobic oxidation of toluene, phenol, and p-cresol by
428 the dissimilatory iron-reducing organism, GS-15. *Appl. Environ. Microbiol.* **1990,** *56*, (6), 1858-
429 1864.
430 (20) Lovley, D. R.; Baedecker, M. J.; Lonergan, D. J.; Cozzarelli, I. M.; Phillips, E. J.; Siegel,
431 D. I., Oxidation of aromatic contaminants coupled to microbial iron reduction. *Nature* **1989,** *339*,
432 (6222), 297-300.

(21) Rotaru, D. E. H.; Franks, A. E.; Orellana, R.; Risso, C.; Nevin, K. P., *Geobacter*: the microbe electric's physiology, ecology, and practical applications. *Adv. Microb. Physiol.* **2011,** *59*, (1).

(22) Podosokorskaya, O. A.; Kadnikov, V. V.; Gavrilov, S. N.; Mardanov, A. V.; Merkel, A. Y.; Karnachuk, O. V.; Ravin, N. V.; Bonch-Osmolovskaya, E. A.; Kublanov, I. V., Characterization of *Melioribacter roseus* gen. nov., sp. nov., a novel facultatively anaerobic thermophilic cellulolytic bacterium from the class *Ignavibacteria*, and a proposal of a novel bacterial phylum *Ignavibacteriae*. *Environ. Microbiol.* **2013,** *15*, (6), 1759-1771.

(23) Iino, T.; Mori, K.; Uchino, Y.; Nakagawa, T.; Harayama, S.; Suzuki, K.-i., *Ignavibacterium album* gen. nov., sp. nov., a moderately thermophilic anaerobic bacterium isolated from microbial mats at a terrestrial hot spring and proposal of *Ignavibacteria* classis nov., for a novel lineage at the periphery of green sulfur bacteria. *Int. J. Syst. Bacteriol* **2010,** *60*, (6), 1376-1382.

(24) Lin, X.; Tfaily, M. M.; Green, S. J.; Steinweg, J. M.; Chanton, P.; Imvittaya, A.; Chanton, J. P.; Cooper, W.; Schadt, C.; Kostka, J. E., Microbial metabolic potential for carbon degradation and nutrient (nitrogen and phosphorus) acquisition in an ombrotrophic peatland. *Appl. Environ. Microbiol.* **2014,** *80*, (11), 3531-3540.

(25) Shi, L.; Richardson, D. J.; Wang, Z.; Kerisit, S. N.; Rosso, K. M.; Zachara, J. M.; Fredrickson, J. K., The roles of outer membrane cytochromes of Shewanella and Geobacter in extracellular electron transfer. *Environ. Microbiol. Rep.* **2009,** *1*, (4), 220-227.

(26) Liu, Y.; Wang, Z.; Liu, J.; Levar, C.; Edwards, M. J.; Babauta, J. T.; Kennedy, D. W.; Shi, Z.; Beyenal, H.; Bond, D. R., A trans-outer membrane porin-cytochrome protein complex

455 for extracellular electron transfer by *Geobacter sulfurreducens* PCA. *Environ. Microbiol. Rep.*
456 **2014,** *6*, (6), 776-785.
457 (27) Reardon, P. N.; Mueller, K. T., Structure of the type IVa major pilin from the electrically
458 conductive bacterial nanowires of Geobacter sulfurreducens. *J. Biol. Chem.* **2013,** *288*, (41),
459 29260-29266.
460 (28) Reguera, G.; McCarthy, K. D.; Mehta, T.; Nicoll, J. S.; Tuominen, M. T.; Lovley, D. R.,
461 Extracellular electron transfer via microbial nanowires. *Nature* **2005,** *435*, (7045), 1098-1101.
462 (29) Wang, X.; Guo, X.; Yang, Y.; Tao, S.; Xing, B., Sorption mechanisms of phenanthrene,
463 lindane, and atrazine with various humic acid fractions from a single soil sample. *Environ. Sci.*
464 *Technol.* **2011,** *45*, (6), 2124-2130.
465 (30) Yang, Y.; Shu, L.; Wang, X.; Xing, B.; Tao, S., Mechanisms regulating bioavailability of
466 phenanthrene sorbed on a peat soil-origin humic substance. *Environ. Tox. Chem.* **2012,** *31*, (7),
467 1431-1437.
468 (31) IHSS, What are Humic Substances? *http://www.humicsubstances.org/whatarehs.html*
469 **2015**.
470 (32) Wagai, R.; Mayer, L. M.; Kitayama, K.; Shirato, Y., Association of organic matter with
471 iron and aluminum across a range of soils determined via selective dissolution techniques
472 coupled with dissolved nitrogen analysis. *Biogeochemistry* **2013,** *112*, (1-3), 95-109.
473 (33) Chang, H.-X.; Yendrek, C. R.; Caetano-Anolles, G.; Hartman, G. L., Genomic
474 characterization of plant cell wall degrading enzymes and in silico analysis of xylanses and
475 polygalacturonases of Fusarium virguliforme. *BMC Microbiol.* **2016,** *16*, (1), 147.

(34) Roden, E. E., Diversion of electron flow from methanogenesis to crystalline Fe (III) oxide reduction in carbon-limited cultures of wetland sediment microorganisms. *Appl. Environ. Microbiol.* **2003,** *69*, (9), 5702-5706.

(35) Roden, E. E.; Urrutia, M. M., Influence of biogenic Fe(II) on bacterial crystalline Fe(III) oxide reduction. *Geomicrobiol. J.* **2002,** *19*, (2), 209-251.

(36) Stookey, L. L., Ferrozine - a new spectrophotometric reagent for iron. *Anal. Chem.* **1970,** *42*, 779-781.

(37) Markowitz, V. M.; Chen, I.-M. A.; Chu, K.; Szeto, E.; Palaniappan, K.; Pillay, M.; Ratner, A.; Huang, J.; Pagani, I.; Tringe, S., IMG/M 4 version of the integrated metagenome comparative analysis system. *Nucl. Acids Res.* **2014,** *42*, (D1), D568-D573.

(38) Wu, Y.-W.; Tang, Y.-H.; Tringe, S. G.; Simmons, B. A.; Singer, S. W., MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome* **2014,** *2*, (1), 26.

(39) Parks, D. H.; Imelfort, M.; Skennerton, C. T.; Hugenholtz, P.; Tyson, G. W., CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **2015,** *25*, (7), 1043-1055.

(40) Darling, A.; Jospin, G.; Lowe, E.; Matsen, F.; Bik, H.; Eisen, J., PhyloSift: phylogenetic analysis of genomes and metagenomes. PeerJ 2: e243. In 2014.

(41) Huson, D. H.; Mitra, S.; Ruscheweyh, H.-J.; Weber, N.; Schuster, S. C., Integrative analysis of environmental sequences using MEGAN4. *Genome Res.* **2011,** *21*, (9), 1552-1560.

(42) Cuskin, F.; Baslé, A.; Ladevèze, S.; Day, A. M.; Gilbert, H. J.; Davies, G. J.; Potocki-Véronèse, G.; Lowe, E. C., The GH130 family of mannoside phosphorylases contains glycoside

498 hydrolases that target β-1, 2-mannosidic linkages in Candida mannan. *J. Biol. Chem.* **2015,** *290*,
499 (41), 25023-25033.
500 (43) Yin, Y.; Mao, X.; Yang, J.; Chen, X.; Mao, F.; Xu, Y., dbCAN: a web resource for
501 automated carbohydrate-active enzyme annotation. *Nucl. Acids Res.* **2012,** *40*, (W1), W445-
502 W451.
503 (44) Kanehisa, M.; Goto, S.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M., Data,
504 information, knowledge and principle: back to metabolism in KEGG. *Nucl. Acids Res.* **2013,** *42*,
505 (D1), D199-D205.
506 (45) Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M., KEGG as a
507 reference resource for gene and protein annotation. *Nucl. Acids Res.* **2016,** *44*, (D1), D457-D462.
508 (46) Kanehisa, M.; Goto, S., KEGG: kyoto encyclopedia of genes and genomes. *Nucl. Acids*
509 *Res.* **2000,** *28*, (1), 27-30.
510 (47) Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K., KEGG: new
511 perspectives on genomes, pathways, diseases and drugs. *Nucl. Acids Res.* **2017,** *45*, (D1), D353-
512 D361.
513 (48) Eddy, S. R. In *A new generation of homology search tools based on probabilistic*
514 *inference*, Genome Info., 2009; 2009; pp 205-211.
515 (49) Finn, R. D.; Clements, J.; Eddy, S. R., HMMER web server: interactive sequence
516 similarity searching. *Nucl. Acids Res.* **2011**, gkr367.
517 (50) Bagos, P. G.; Liakopoulos, T. D.; Spyropoulos, I. C.; Hamodrakas, S. J., PRED-TMBB: a
518 web server for predicting the topology of β-barrel outer membrane proteins. *Nucl. Acids Res.*
519 **2004,** *32*, (suppl_2), W400-W404.

(51) Markowitz, V. M.; Chen, I.-M. A.; Palaniappan, K.; Chu, K.; Szeto, E.; Grechkin, Y.; Ratner, A.; Jacob, B.; Huang, J.; Williams, P., IMG: the integrated microbial genomes database and comparative analysis system. *Nucl. Acids Res.* **2011,** *40*, (D1), D115-D122.

(52) Reguera, G.; Pollina, R. B.; Nicoll, J. S.; Lovley, D. R., Possible nonconductive role of *Geobacter sulfurreducens* pilus nanowires in biofilm formation. *J. Bacteriol.* **2007,** *189*, (5), 2125-2127.

(53) Pie, J.; Kim, B.; Grishin, N., PROMALS3D: a tool for multiple sequence and structure alignment. *Nucl. Acids Res.* **2008,** *36*, (7), 2295-300.

(54) Melton, E. D.; Swanner, E. D.; Behrens, S.; Schmidt, C.; Kappler, A., The interplay of microbially mediated and abiotic reactions in the biogeochemical Fe cycle. *Nat. Rev. Microbiol.* **2014,** *12*, (12), 797-808.

(55) Nemergut, D. R.; Cleveland, C. C.; Wieder, W. R.; Washenberger, C. L.; Townsend, A. R., Plot-scale manipulations of organic matter inputs to soils correlate with shifts in microbial community composition in a lowland tropical rain forest. *Soil Biol. Biochem.* **2010,** *42*, (12), 2153-2160.

(56) Coolen, M. J.; van de Giessen, J.; Zhu, E. Y.; Wuchter, C., Bioavailability of soil organic matter and microbial community dynamics upon permafrost thaw. *Environ. Microbiol.* **2011,** *13*, (8), 2299-2314.

(57) Mhuantong, W.; Charoensawan, V.; Kanokratana, P.; Tangphatsornruang, S.; Champreda, V., Comparative analysis of sugarcane bagasse metagenome reveals unique and conserved biomass-degrading enzymes among lignocellulolytic microbial communities. *Biotechnol. Fuels* **2015,** *8*, (1), 16.

(58) Colombo, C.; Palumbo, G.; He, J.-Z.; Pinton, R.; Cesco, S., Review on iron availability in soil: interaction of Fe minerals, plants, and microbes. *J. Soil and Sed.* **2014,** *14*, (3), 538-548.

(59) Fortney, N. W.; He, S.; Converse, B.; Beard, B.; Johnson, C.; Boyd, E. S.; Roden, E., Microbial Fe(III) oxide reduction potential in Chocolate Pots hot spring, Yellowstone National Park. *Geobiology* **2016,** *14*, 255-275.

(60) Coates, J.; Ellis, D.; Lovley, D., *Geothrix fermentans* gen. nov. sp. nov., an acetate-oxidizing Fe(III) reducer capable of growth via fermentation. *Int. J. Syst. Bacteriol* **1999,** *49*, 1615-1622.

(61) Warnecke, F.; Luginbühl, P.; Ivanova, N.; Ghassemian, M.; Richardson, T. H.; Stege, J. T.; Cayouette, M.; McHardy, A. C.; Djordjevic, G.; Aboushadi, N., Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **2007,** *450*, (7169), 560.

(62) Liu, N.; Zhang, L.; Zhou, H.; Zhang, M.; Yan, X.; Wang, Q.; Long, Y.; Xie, L.; Wang, S.; Huang, Y., Metagenomic insights into metabolic capacities of the gut microbiota in a fungus-cultivating termite (Odontotermes yunnanensis). *PLoS One* **2013,** *8*, (7), e69184.

(63) He, S.; Ivanova, N.; Kirton, E.; Allgaier, M.; Bergin, C.; Scheffrahn, R. H.; Kyrpides, N. C.; Warnecke, F.; Tringe, S. G.; Hugenholtz, P., Comparative metagenomic and metatranscriptomic analysis of hindgut paunch microbiota in wood-and dung-feeding higher termites. *PLoS One* **2013,** *8*, (4), e61126.

(64) Wang, C.; Dong, D.; Wang, H.; Müller, K.; Qin, Y.; Wang, H.; Wu, W., Metagenomic analysis of microbial consortia enriched from compost: new insights into the role of Actinobacteria in lignocellulose decomposition. *Biotechnol. Fuels* **2016,** *9*, (1), 22.

563 (65) Howe, A.; Yang, F.; Williams, R. J.; Meyer, F.; Hofmockel, K. S., Identification of the
564 Core Set of Carbon-Associated Genes in a Bioenergy Grassland Soil. *PLoS One* **2016,** *11*, (11),
565 e0166578.
566 (66) Schellenberger, S.; Kolb, S.; Drake, H. L., Metabolic responses of novel cellulolytic and
567 saccharolytic agricultural soil bacteria to oxygen. *Environ. Microbiol.* **2010,** *12*, (4), 845-861.
568 (67) Munir, R. I.; Schellenberg, J.; Henrissat, B.; Verbeke, T. J.; Sparling, R.; Levin, D. B.,
569 Comparative analysis of carbohydrate active enzymes in Clostridium termitidis CT1112 reveals
570 complex carbohydrate degradation ability. *PLoS One* **2014,** *9*, (8), e104260.
571 (68) van Munster, J. M.; Daly, P.; Delmas, S.; Pullan, S. T.; Blythe, M. J.; Malla, S.;
572 Kokolski, M.; Noltorp, E. C.; Wennberg, K.; Fetherston, R., The role of carbon starvation in the
573 induction of enzymes that degrade plant-derived carbohydrates in *Aspergillus niger*. *Fungal Gen.*
574 *Biol.* **2014,** *72*, 34-47.
575 (69) Adhikari, D.; Poulson, S. R.; Sumaila, S.; Dynes, J. J.; McBeth, J. M.; Yang, Y.,
576 Asynchronous reductive release of iron and organic carbon from hematite-humic acid
577 complexes. *Chem. Geol.* **2016,** *430*, 13-20.
578 (70) Kalbitz, K.; Schmerwitz, J.; Schwesig, D.; Matzner, E., Biodegradation of soil-derived
579 dissolved organic matter as related to its properties. *Geoderma* **2003,** *113*, (3), 273-291.
580 (71) Ueno, A.; Shimizu, S.; Tamamura, S.; Okuyama, H.; Naganuma, T.; Kaneko, K.,
581 Anaerobic decomposition of humic substances by Clostridium from the deep subsurface. *Sci.*
582 *Rep.* **2016,** *6*.
583 (72) Yang, Y.; Shu, L.; Wang, X.; Xing, B.; Tao, S., Impact of de-ashing humic acid and
584 humin on organic matter structural properties and sorption mechanisms of phenanthrene.
585 *Environ. Sci. Technol.* **2011,** *45*, (9), 3996-4002.

586 (73) Rabenhorst, M. C.; Swanson, D., Histosols. In *Handbook of Soil Science*, Summer, M. E.,
587 Ed. CRC Press: Boca Raton, 1999.
588 (74) Yang, Y., University of Nevada-Reno, Reno, NV, Unpublished work, 2018.
589 (75) Wagai, R.; Mayer, L. M., Sorptive stabilization of organic matter in soils by hydrous iron
590 oxides. *Geochim. Cosmochim. Acta* **2007,** *71*, (1), 25-35.
591 (76) Lalonde, K.; Mucci, A.; Ouellet, A.; Gélinas, Y., Preservation of organic matter in
592 sediments promoted by iron. *Nature* **2012,** *483*, (7388), 198-200.
593
594

Supporting Information

Dual role of humic substances as electron donor and shuttle for dissimilatory iron reduction

Noah Stern[1], Jacqueline Mejia[1], Shaomei He[2,3], Yu Yang[4], Matthew Ginder-Vogel[1], and Eric E Roden*,[2]

[1] Department of Civil and Environmental Engineering, Environmental Chemistry and Technology Program, University of Wisconsin-Madison, Madison, WI 53706, USA

[2] Department of Geoscience, University of Wisconsin-Madison, Madison, WI 53706, USA

[3] Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53706, USA

[4] Department of Civil and Environmental Engineering, University of Nevada-Reno, Reno, NV, 89557, USA

*Corresponding author. Email: eroden@geology.wisc.edu. Phone: 608-890-7024. Fax: 608-262-0693.

The Supporting Information contains 3 Supporting Text sections, 4 Figures, and 12 Tables that complement the metagenomic results presented in the manuscript, including a description of the porin-cytochrome complex (PCC) system found in *Geothrix fermentans* (Supplemental Text 1 and Figure S1); results of the PilA gene search (Figure S2); detailed information about the assembly of raw reads using CLC Genomics workbench (Table S1); a list of the phylogenetic marker genes (Table S2) used to determine the microbial community composition (Table S3); an illustration of the presence of carbohydrate active enzyme (CAZyme) modules and their taxonomic assignments (Figures S3 and S4); the abundance and known activity of CAZyme families and KEGG aromatic degradation genes that had an abundance at least two times higher in the presence of HA or HM compared to the glucose only treatments (Supporting Text 2 and Tables S4-S7); a description and results of screening PCC-containing draft genomes for KEGG aromatic degradation pathways (Supporting Text 3 and Tables S8-S12); and calculations to estimates the abundance of metabolizable HA in a representative histosol.

Supporting Text 1- *Geothrix fermentans* sp. PCC Description. The PCC gene system was identified in *Geothrix fermentans* sp. nov., by hidden Markov model designed to target the beta-barrel porin protein (OmbB/OmbC) from *Geobacter sulfurreducens* PCA (details in Materials and Methods) (Figure S 3). Protein analysis for cellular location, beta-barrel porin trans-membrane motifs and c-type cytochrome heme-binding cites revealed that the PCC in *Geothrix fermentans* sp. nov., contains two periplasmic c-type cytochromes labeled OmaB1 and OmaB2 containing 5 and 6 heme-binding cites respectively (Figure S 3). The analogous OmcB extracellular c-type cytochrome contains 10 heme-binding cites. It is worth noting that the total number of heme-binding sites in the *Geobacter sulfurreducens* PCA OmaB-OmcB gene cluster is 20 and the total number of heme-binding sites in the *Geothrix fermentans* sp. nov., OmaB1-OmaB2-OmcB gene cluster is 21. Comparison between the PCC of *Geothrix fermentans* sp. nov., found in this experiment and the previously published whole genome for *Geothrix fermentans* DSM 14018[2] showed a matching PCC gene system with three additional extracellular decaheme c-type cytochromes.

The DIR capabilities of *Geothrix fermentans* are well understood[3,4] however the EET gene system utilized to move electrons from inside the cell to electron acceptors outside the cell has not been documented. The *Geothrix fermentans* sp. nov., from this study was identified to contain the PCC first discovered in *Geobacter sulfurreducens PCA* and later found in organisms such as *Anaeromyxobacter dehalogenans* 2CP-1, *Denitrovibrio acetiphilus* DSM 12809, *Ignavibacterium album* JCM 16511 and *Desulfuromonas acetoxidans* DSM 684.[5] Although this study does provide evidence that *Geothrix fermentans* sp. nov., contains the necessary genes for a complete PCC metabolism pathway further studies are needed to provide proof of this pathways functionality.

Supporting Text 2 - Key Enzymatic Systems Involved in Humic Substance Metabolism. CAZyme families and KEGG aromatic degradation genes with abundance values at least two times higher in metagenomes amended with humic substances (humic acids-HA or humins-HM) compared to glucose only treatments are illustrated in Tables S4, S5, S6 and S7. The known CAZyme classification was determined by previous studies,[6-18] and is not specified for all CAZy families.

Supporting Text 3 – Aromatic Carbon Degradation Analysis. Genome reconstruction (i.e. binning) from metagenomes was carried out by a coverage-composition algorithm using MaxBin.[19] The completeness and potential contamination of the bins was assessed using CheckM.[20] The curated high quality bins (i.e., draft genomes) were searched for PCC genes, which resulted in classifications to as *Geobacter*, *Goethrix*, *Ignavibacteriales*, *Anaeromyxobacter* and *Prolixibacter*. These draft genomes and, for comparison genomes downloaded from IMG, were then screened for KEGG genes involved in aromatic degradation pathways (Tables S8, S9, S10 and S11).

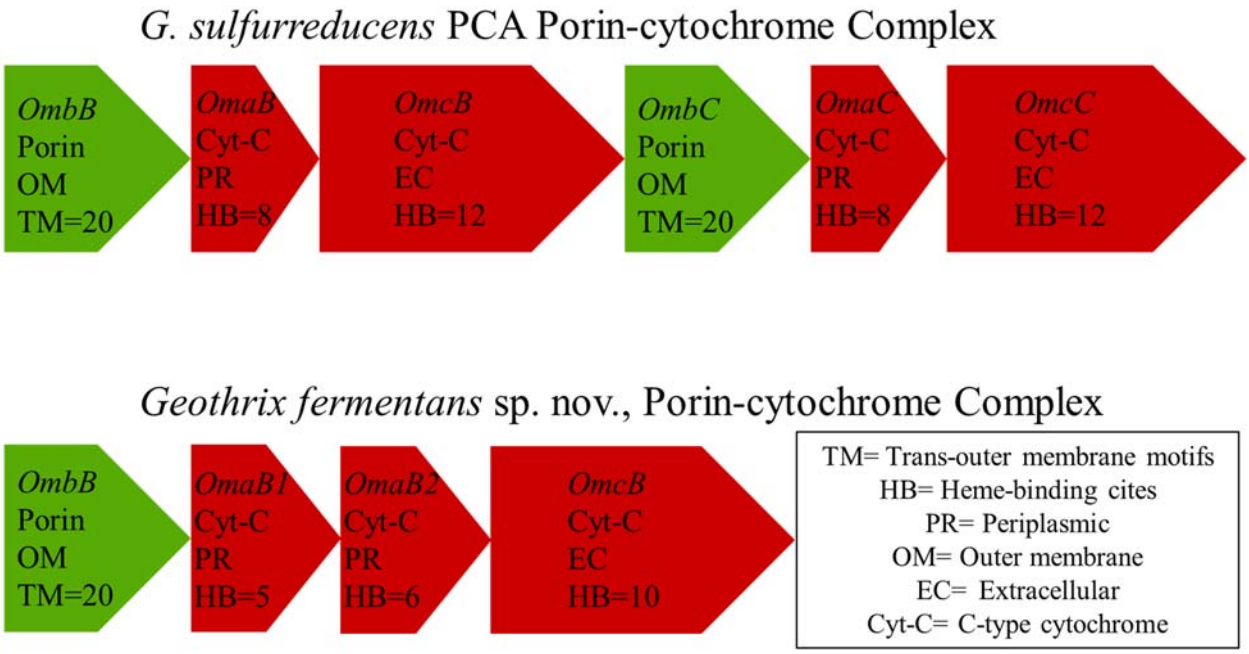


Figure S1. Comparison of PCC in *Geobacter sulfurreducens* PCA and *Geothrix fermentans* sp. nov., identified from the metagenomes of this study.[1]

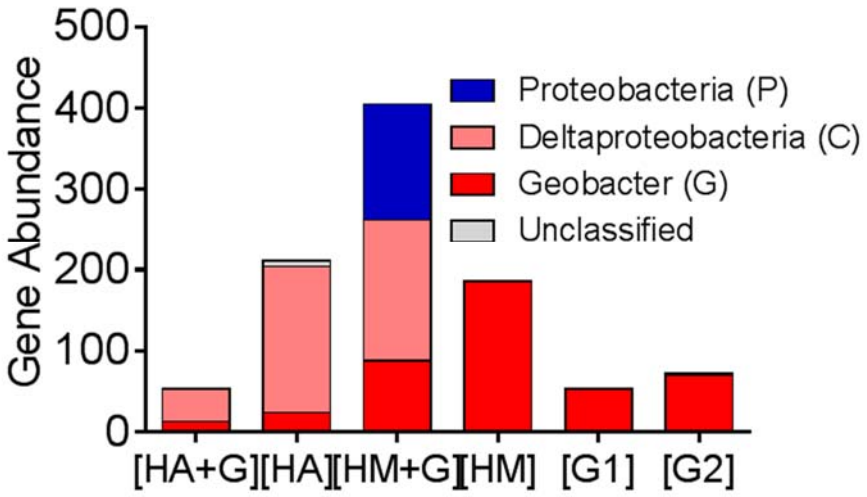


Figure S2. PilA gene abundance for the HA and HM experiments. Assigned gene copies reflect the hits multiplied by the IMG scaffold read depth and normalized to the total base pairs in each metagenome. Abbreviations: HA = Humic Acid; G = Glucose; HM = Humin. G1 and G2 represent replicate glucose-only cultures.

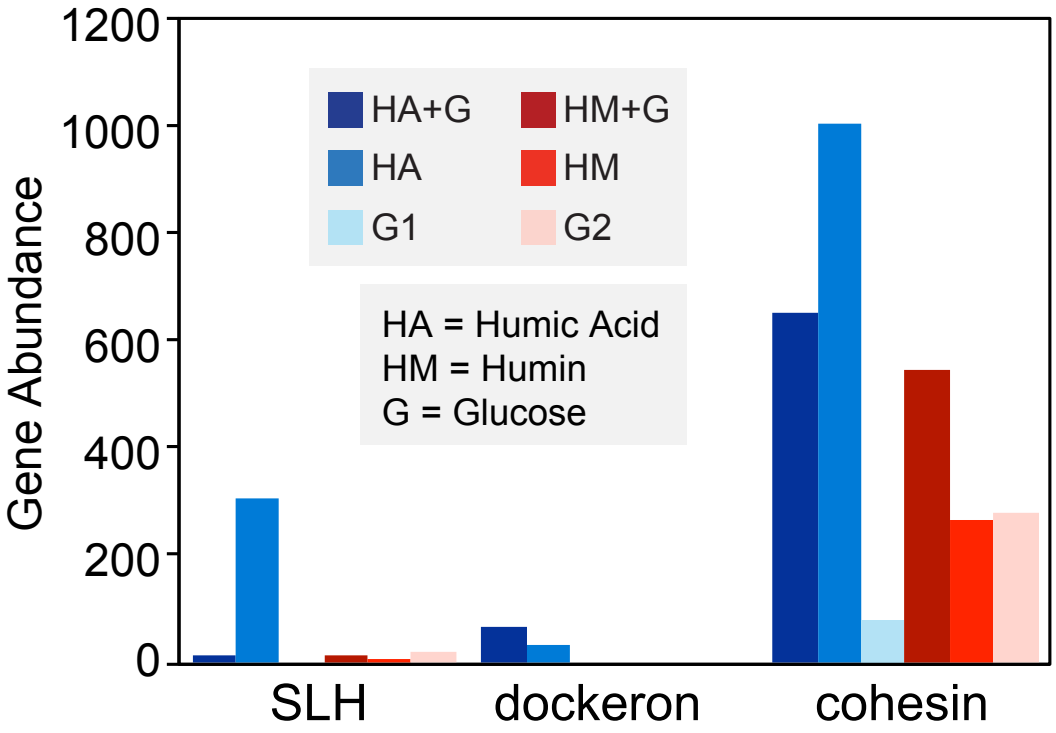


Figure S3. Abundance of genes categorized within a carbohydrate active enzyme (CAZyme) module (SLH, docketin and cohesin), in the HA and HM data sets. The gene abundance was normalized by the total assembled base pairs.

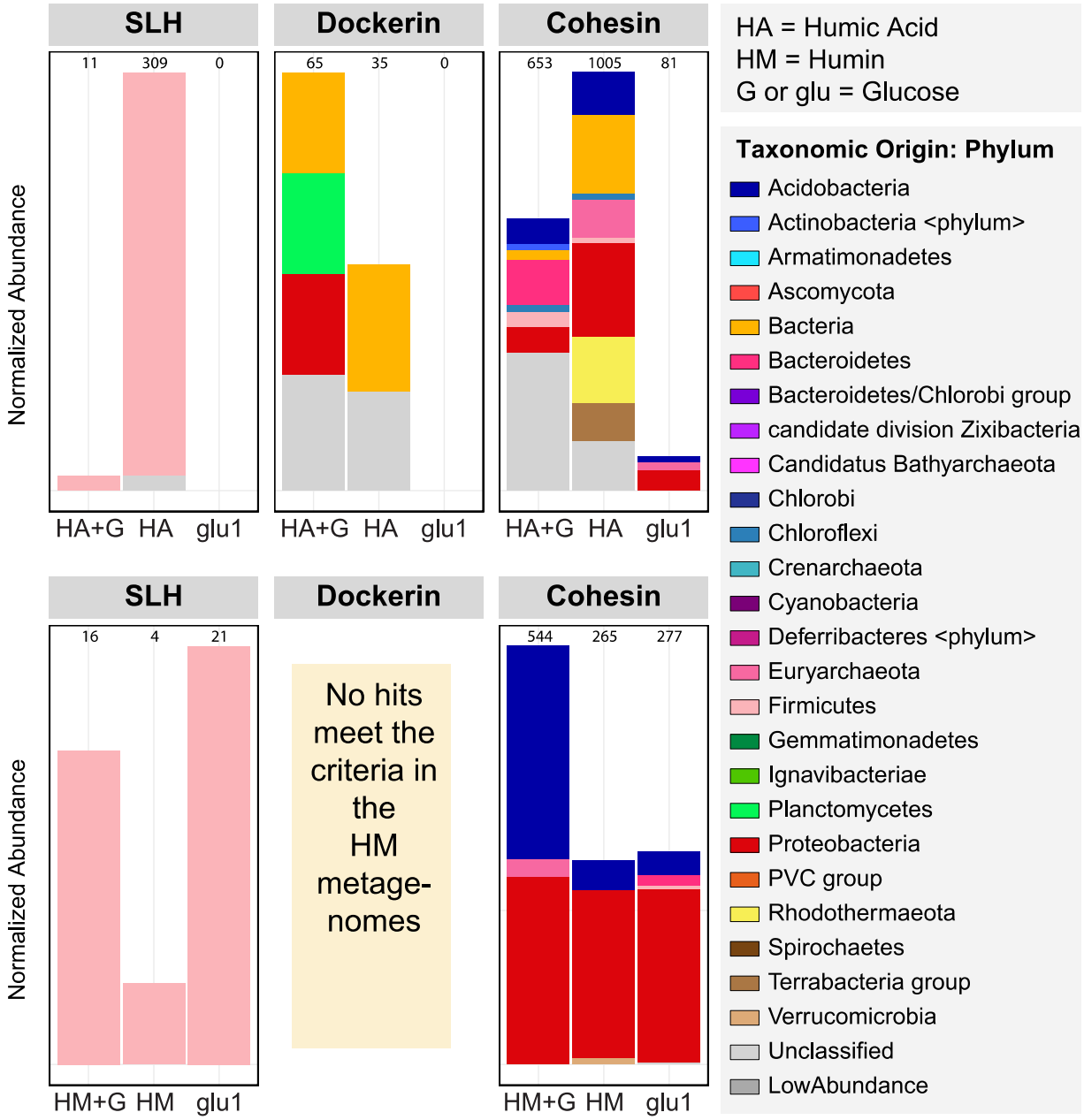


Figure S4. Taxonomic origin of CAZy modules present in all treatments. The y-axis represents the relative abundance of the phyla and the number at the top of the bar represents the total abundance of a specific modules and for a specific treatment.

Table S1. Metagenomes assembly with CLC Genomic Workbench information.

Metagenome Name	HA1 - Humic Acid and Glucose	HA2 - Glucose Only	HA3 - Humic Acid Only	HM1 - Humin and Glucose	HM2 - Glucose Only	HM3 - Humin Only
# raw reads	104473306	91226740	96492596	101724132	103934230	113463654
# reads after quality trimming	103265407	90181765	95508545	96399945	97833792	106864039
Avg length of reads matched after trimming	100.09	100.05	100.09	98.57	98.41	98.42
Total bases of reads matched after trimming	10335773099	9022764233	9559836211	9502323849	9628120889	10517032546
# reads matched after assembly	69236438	77944508	59796089	81074555	85566986	83953644
Avg length of reads matched after assembly	100.17	100.11	100.21	98.58	98.41	98.41
Total bases of reads matched after assembly*	6935661698	7802952571	5991905329	7992678301	8420617307	8261511278
# reads NOT matched after assembly	34028969	12237257	35712456	15325390	12266806	22910395
Avg length of reads NOT matched after assembly	99.92	99.68	99.91	98.51	98.44	98.45
Total bases of reads NOT matched after assembly	3400111401	1219811662	3567930882	1509645548	1207503582	2255521268
% raw reads assembled	67%	86%	63%	84%	87%	79%
Contig N75:	618	753	531	876	989	732
Contig N50:	1924	6160	1605	5870	5563	2717
Contig N25:	13420	65123	9535	82088	40778	25972
Contig Minimum length:	119	79	68	92	106	99
Contig Maximum length:	1042875	1235177	1262335	1804791	1786503	1581748
Contig Average length:	1010	1224	908	1312	1383	1110
Contig Total length	306895909	171117903	340747055	208247969	193292646	175066219
# Contigs	303938	139850	375441	158677	139783	157672

* This value was used to normalize the coverage among the different samples

Table S2. 32 COGs representing phylogenetic marker chosen for microbial community analysis.

COG ID	COG Name	Consensus COG length (amino acids)
COG0016	Phenylalanyl-tRNA synthetase alpha subunit	335
COG0072	Phenylalanyl-tRNA synthetase beta subunit	650
COG0081	Ribosomal protein L1	228
COG0244	Ribosomal protein L10	175
COG0080	Ribosomal protein L11	141
COG0102	Ribosomal protein L13	148
COG0093	Ribosomal protein L14	122
COG0200	Ribosomal protein L15	152
COG0197	Ribosomal protein L16/L10AE	146
COG0256	Ribosomal protein L18	125
COG0090	Ribosomal protein L2	275
COG0091	Ribosomal protein L22	120
COG0089	Ribosomal protein L23	94
COG0198	Ribosomal protein L24	104
COG0087	Ribosomal protein L3	218
COG0088	Ribosomal protein L4	214
COG0094	Ribosomal protein L5	180
COG0097	Ribosomal protein L6P/L9E	178
COG0051	Ribosomal protein S10	104
COG0100	Ribosomal protein S11	129
COG0048	Ribosomal protein S12	129
COG0099	Ribosomal protein S13	121
COG0184	Ribosomal protein S15P/S13E	89
COG0186	Ribosomal protein S17	87
COG0185	Ribosomal protein S19	93
COG0052	Ribosomal protein S2	252
COG0092	Ribosomal protein S3	233
COG0098	Ribosomal protein S5	181
COG0049	Ribosomal protein S7	148
COG0096	Ribosomal protein S8	132
COG0103	Ribosomal protein S9	130
COG0532	Translation initiation factor IF-2, a GTPase	509

Table S3. Microbial community composition based on phylogenetic marker genes found in the metagenomes. Taxonomic assignments of these genes was carried out by BLASTP using the Ref_Seq protein database.

Phylogenetic Markers by MEGAN_RefSeq_COG len normalized						
Phylum	Humic Acid + Glucose	Humic Acid Only	Glucose Only 1	Humin + Glucose	Humin Only	Glucose Only 2
Spirochaetes	20.8	5.2	0.0	0.0	0.0	0.1
Proteobacteria	13.7	23.5	84.2	70.9	86.5	73.2
Euryarchaeota	11.6	18.1	1.0	1.6	0.0	0.0
Bacteroidetes	11.2	5.5	0.9	3.5	0.0	0.3
Ignavibacteriae	10.1	10.5	0.0	0.0	0.0	0.0
Acidobacteria	8.9	7.3	11.7	22.4	9.1	24.3
Firmicutes	4.7	9.3	1.1	0.1	0.0	0.2
Chloroflexi	2.2	7.3	0.0	0.0	0.0	0.0
Lentisphaerae	1.8	0.0	0.0	0.0	0.0	0.0
Actinobacteria	2.1	5.8	0.0	0.0	0.0	0.0
Verrucomicrobia	1.7	0.5	0.2	1.2	4.0	0.2
Candidatus Kryptonia	1.3	0.9	0.0	0.0	0.0	0.0
Crenarchaeota	0.5	0.8	0.0	0.0	0.0	0.0
Terrabacteria group	0.4	0.7	0.1	0.0	0.0	0.0
Chrysiogenetes	0.3	0.0	0.0	0.0	0.0	0.0
Thermotogae	0.3	0.0	0.0	0.0	0.0	0.0
FCB group	0.2	0.7	0.0	0.0	0.0	0.0
Deferribacteres	0.1	0.2	0.0	0.0	0.0	0.0
Aquificae	0.1	0.0	0.0	0.0	0.0	0.0
Thaumarchaeota	0.1	0.1	0.0	0.0	0.0	0.0
Nitrospirae	0.1	0.0	0.0	0.0	0.0	0.0
Planctomycetes	0.0	0.6	0.0	0.0	0.0	0.0
Cyanobacteria	0.0	0.0	0.0	0.0	0.0	0.0
Unclassified	8.0	2.8	0.8	0.2	0.3	1.7
Total	100.0	100.0	100.0	100.0	100.0	100.0

Table S4. 129 CAZy families with abundance values at least two times higher in the metagenomes containing humic acid (HA) versus those with only glucose (G).

Glycoside Hydrolases (GHs)					
CAZy Family	Known CAZyme Activities	Known CAZyme Classification	HA + G	HA	G1
GH74	endoglucanase; oligoxyloglucan reducing end-specific cellobiohydrolase; xyloglucanase	Cellulases and endohemicellulases	15960	14280	1160
GH39	α-L-iduronidase; β-xylosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	7385	619	78
GH2	β-galactosidase; β-mannosidase; β-glucuronidase; α-L-arabinofuranosidase; mannosylglycoprotein endo-β-mannosidase; exo-β-glucosaminidase	Debranching enzymes and Oligosaccharide-degrading enzymes	6228	3316	1473
GH5	endo-β-1,4-glucanase / cellulase; endo-β-1,4-xylanase; β-glucosidase; β-mannosidase; β-glucosylceramidase; glucan β-1,3-glucosidase; licheninase; exo-β-1,4-glucanase / cellodextrinase; glucan endo-1,6-β-glucosidase; mannan endo-β-1,4-mannosidase; cellulose β-1,4-cellobiosidase; steryl β-glucosidase; endoglycoceramidase; chitosanase; β-primeverosidase; xyloglucan-specific endo-β-1,4-glucanase; endo-β-1,6-galactanase; hesperidin 6-O-α-L-rhamnosyl-β-glucosidase; β-1,3-mannanase; arabinoxylan-specific endo-β-1,4-xylanase; mannan transglycosylase	Cellulases and endohemicellulases	6150	2394	957
GH33	sialidase or neuraminidase; trans-sialidase; 2-keto-3-deoxynononic acid hydrolase; anhydrosialidase; 3-deoxy-D-manno-octulosonic-acid hydrolase	Debranching enzymes and Oligosaccharide-degrading enzymes	4834	3301	649
GH78	α-L-rhamnosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	4824	2682	69
GH42	β-galactosidase; α-L-arabinopyranosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	4622	498	40
GH29	α-L-fucosidase; α-1,3/1,4-L-fucosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	4385	2311	87

GH130	β-1,4-mannosylglucose phosphorylase; β-1,4-mannooligosaccharide phosphorylase; β-1,4-mannosyl-N-acetyl-glucosamine phosphorylase; β-1,2-mannobiose phosphorylase; β-1,2-oligomannan phosphorylase; β-1,2-mannosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	3896	2353	419
GH28	polygalacturonase; exo-polygalacturonase; exo-polygalacturonosidase; rhamnogalacturonase; rhamnogalacturonan α-1,2-galacturonohydrolase; rhamnogalacturonan α-L-rhamnopyranohydrolase; endo-xylogalacturonan hydrolase	Pectinases	2494	1644	134
GH127	β-L-arabinofuranosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	2427	1947	470
GH110	α-galactosidase; α-1,3-galactosidase	NA	2403	233	5
GH43	β-xylosidase; α-L-arabinofuranosidase; arabinanase; xylanase; galactan 1,3-β-galactosidase; α-1,2-L-arabinofuranosidase; exo-α-1,5-L-arabinofuranosidase; [inverting] exo-α-1,5-L-arabinanase; β-1,3-xylosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	2381	1850	578
GH92	mannosyl-oligosaccharide α-1,2-mannosidase; mannosyl-oligosaccharide α-1,3-mannosidase; mannosyl-oligosaccharide α-1,6-mannosidase; α-mannosidase; α-1,2-mannosidase; α-1,3-mannosidase; α-1,4-mannosidase; mannosyl-1-phosphodiester α-1,P-mannosidase	CAZy families processing fungal cell wall	2239	1723	86
GH20	β-hexosaminidase; lacto-N-biosidase; β-1,6-N-acetylglucosaminidase); β-6-SO3-N-acetylglucosaminidase	CAZy families processing fungal cell wall	2236	1235	141
GH38	α-mannosidase; mannosyl-oligosaccharide α-1,2-mannosidase; mannosyl-oligosaccharide α-1,3-1,6-mannosidase; α-2-O-mannosylglycerate hydrolase; mannosyl-oligosaccharide α-1,3-mannosidase	CAZy families processing fungal cell wall	2225	1352	513
GH57	α-amylase; α-galactosidase; amylopullulanase; cyclomaltodextrinase; branching enzyme; 4-α-glucanotransferase	Debranching enzymes and Oligosaccharide-degrading enzymes	2073	2001	440
GH51	endoglucanase; endo-β-1,4-xylanase; β-xylosidase; α-L-arabinofuranosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	1838	1011	84

GH36	α-galactosidase; α-N-acetylgalactosaminidase; stachyose synthase; raffinose synthase	Debranching enzymes and Oligosaccharide-degrading enzymes	1798	616	70
GH106	α-L-rhamnosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	1716	908	18
GH105	unsaturated rhamnogalacturonyl hydrolase; d-4,5-unsaturated β-glucuronyl hydrolase	Pectinases	1655	986	380
GH95	α-L-fucosidase; α-1,2-L-fucosidase; α-L-galactosidase	Cellulases and endohemicellulases	1473	997	43
GH16	xyloglucan:xyloglucosyltransferase; keratan-sulfate endo-1,4-β-galactosidase; endo-1,3-β-glucanase; endo-1,3(4)-β-glucanase; licheninase; β-agarase; κ-carrageenase; xyloglucanase; endo-β-1,3-galactanase; β-porphyranase; hyaluronidase; endo-β-1,4-galactosidase; chitin β-1,6-glucanosyltransferase; endo-β-1,4-galactosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	1422	1400	451
GH50	β-agarase	NA	1372	470	30
GH99	glycoprotein endo-α-1,2-mannosidase; mannan endo-1,2-α-mannanase	CAZy families processing fungal cell wall	1066	216	19
GH35	β-galactosidase; exo-β-glucosaminidase; exo-β-1,4-galactanase; β-1,3-galactosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	941	462	92
GH30	endo-β-1,4-xylanase; β-glucosidase; β-glucuronidase; β-xylosidase; β-fucosidase; glucosylceramidase; β-1,6-glucanase; glucuronoarabinoxylan endo-β-1,4-xylanase; endo-β-1,6-galactanase; [reducing end] β-xylosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	842	795	308
GH116	β-glucosidase; β-xylosidase; acid β-glucosidase/β-glucosylceramidase; β-N-acetylglucosaminidase	Debranching enzymes and Oligosaccharide-degrading enzymes	819	639	42
GH27	α-galactosidase; α-N-acetylgalactosaminidase; isomalto-dextranase; β-L-arabinopyranosidase; galactan:galactan galactosyltransferase	Debranching enzymes and Oligosaccharide-degrading enzymes	799	570	37
GH53	endo-β-1,4-galactanase	Debranching enzymes and Oligosaccharide-degrading enzymes	766	285	23
GH10	endo-1,4-β-xylanase; endo-1,3-β-xylanase; tomatinase; xylan endotransglycosylase	Cellulases and endohemicellulases	750	763	55

GH9	endoglucanase; endo-β-1,3(4)-glucanase / lichenase-laminarinase; β-glucosidase; lichenase / endo-β-1,3-1,4-glucanase; exo-β-1,4-glucanase / cellodextrinase; cellobiohydrolase; xyloglucan-specific endo-β-1,4-glucanase / endo-xyloglucanase; exo-β-glucosaminidase	Cellulases and endohemicellulases	646	743	94
GH93	exo-α-L-1,5-arabinanase	Debranching enzymes and Oligosaccharide-degrading enzymes	613	853	0
GH71	α-1,3-glucanase	Debranching enzymes and Oligosaccharide-degrading enzymes	528	13	0
GH108	N-acetylmuramidase	NA	515	1153	229
GH129	α-N-acetylgalactosaminidase	NA	511	21	0
GH97	glucoamylase; α-glucosidase; α-galactosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	496	453	66
GH123	β-N-acetylgalactosaminidase; glycosphingolipid β-N-acetylgalactosaminidase	NA	491	254	0
GH12	endoglucanase; xyloglucan hydrolase; β-1,3-1,4-glucanase; xyloglucan endotransglycosylase	Cellulases and endohemicellulases	473	204	4
GH117	α-1,3-L-neoagarooligosaccharide hydrolase; α-1,3-L-neoagarobiase/neoagarobiose hydrolase	NA	453	337	4
GH26	β-mannanase; exo-β-1,4-mannobiohydrolase; β-1,3-xylanase; lichenase / endo-β-1,3-1,4-glucanase; mannobiose-producing exo-β-mannanase	Cellulases and endohemicellulases	410	375	37
GH89	α-N-acetylglucosaminidase	CAZy families processing fungal cell wall	378	310	5
GH44	endoglucanase; xyloglucanase	Cellulases and endohemicellulases	350	189	23
GH125	exo-α-1,6-mannosidase	CAZy families processing fungal cell wall	332	169	0
GH64	β-1,3-glucanase	Debranching enzymes and Oligosaccharide-degrading enzymes	173	217	10
GH25	lysozyme	NA	145	458	63
GH47	α-mannosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	142	168	0
GH17	glucan endo-1,3-β-glucosidase; glucan 1,3-β-glucosidase; licheninase; ABA-specific β-glucosidase; β-1,3-glucanosyltransglycosylase	Cell wall elongation	137	86	8
GH100	alkaline and neutral invertase	Disaccharide Degrading	129	223	19

GH76	α-1,6-mannanase	CAZy families processing fungal cell wall	105	164	17
GH120	β-xylosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	93	164	25
GH128	β-1,3-glucanase	CAZy families processing fungal cell wall	50	59	0
GH66	cycloisomaltooligosaccharide glucanotransferase; dextranase	NA	48	231	24
GH55	exo-β-1,3-glucanase; endo-β-1,3-glucanase	CAZy families processing fungal cell wall	46	148	0
GH62	α-L-arabinofuranosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	38	55	6
GH11	endo-β-1,4-xylanase; endo-β-1,3-xylanase	Cellulases and endohemicellulases	36	199	0
GH54	α-L-arabinofuranosidase; β-xylosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	35	100	0
GH67	α-glucuronidase; xylan α-1,2-glucuronidase	Debranching enzymes and Oligosaccharide-degrading enzymes	23	35	0
GH79	β-glucuronidase; hyaluronoglucuronidase; heparanase; baicalin β-glucuronidase; β-4-O-methyl-glucuronidase	Pectinases	22	16	0
GH59	β-galactosidase; galactocerebrosidase	Cellulases and endohemicellulases	18	55	0
GH84	N-acetyl β-glucosaminidase; hyaluronidase; [protein]-3-O-(GlcNAc)-L-Ser/Thr β-N-acetylglucosaminidase	NA	12	52	0
Glycosyl Transferases (GTs)					
CAZy Family	Known CAZyme Activities		HA + G	HA	G1
GT5	UDP-Glc: glycogen glucosyltransferase; ADP-Glc: starch glucosyltransferase; NDP-Glc: starch glucosyltransferase; UDP-Glc: α-1,3-glucan synthase; UDP-Glc: α-1,4-glucan synthase		2376	2233	1102
GT39	Dol-P-Man: protein α-mannosyltransferase		1051	799	335
GT66	dolichyl-diphosphooligosaccharide—protein glycotransferase; undecaprenyl-diphosphooligosaccharide—protein glycotransferase		718	1731	257
GT81	NDP-Glc: glucosyl-3-phosphoglycerate synthase; NDP-Man: mannosyl-3-phosphoglycerate synthase; ADP-Glc: glucosyl-2-glycerate synthase		461	884	80
GT3	glycogen synthase		249	400	37
GT94	GDP-Man: GlcA-β-1,2-Man-α-1,3-Glc-β-1,4-Glc-α-1-PP-undecaprenol β-1,4-mannosyltransferase		167	371	42
GT10	galactoside α-1,3/1,4-L-fucosyltransferase; galactoside α-1,3-L-fucosyltransferase; glycoprotein α-1,3-L-fucosyltransferase		145	127	0
GT75	UDP-Glc: self-glucosylating β-glucosyltransferase; UDP-L-arabinopyranose mutase		103	134	0
GT89	β-D-arabinofuranosyl-1-monophosphoryldecaprenol : arabinan β-1,2-arabinofuranosyltransferase		92	406	37

GT74	α-1,2-L-fucosyltransferase	88	71	9
GT13	α-1,3-mannosyl-glycoprotein β-1,2-N-acetylglucosaminyltransferase	64	42	0
GT92	UDP-Gal: N-glycan core α-1,6-fucoside β-1,4-galactosyltransferase; UDP-Gal: β-galactoside β-1,4-galactosyltransferase	40	28	0
GT77	α-xylosyltransferase; α-1,3-galactosyltransferase; arabinosyltransferase; arabinosyltransferase	15	11	0
GT33	GDP-Man: chitobiosyldiphosphodolichol β-mannosyltransferase	7	180	0
Carbon Binding Modules (CBMs)				
CAZy Family	Known CAZyme Activities	HA + G	HA	G1
CBM9	Binds to cellulose	10344	3130	849
CBM44	Binds to cellulose and xyloglucan	5329	9754	1808
CBM32	Binds to galactose, lactose, polygalacturonic acid and LacNAc (β-D-galactosyl-1,4-β-D-N-acetylglucosamine)	4933	2900	613
CBM67	Binds to L-rhamnose	3918	1841	208
CBM16	Binds to cellulose and glucomannan	2437	757	187
CBM66	Binds to the terminal fructoside residue of fructans	1815	2373	658
CBM4	Binds to xylan, β-1,3-glucan, β-1,3-1,4-glucan, β-1,6-glucan and amorphous cellulose but not with crystalline cellulose.	1653	447	67
CBM51	Attached to various enzymes from families GH2, GH27, GH31, GH95, GH98 and GH101. Binds to galactose and to blood group A/B-antigens	1076	443	30
CBM40	Attached to various enzymes from GH 33 family. Binds to sialic acid	609	448	71
CBM57	binds to various glycosidases	567	987	28
CBM35	Binds to xylan and β-galactan, and decorated soluble mannans and mannooligosaccharides	462	223	6
CBM20	Binds to granular starch and interacts strongly with cyclodextrins	405	653	124
CBM22	Binds to xylan and has affinity with mixed β-1,3/β-1,4-glucans	346	71	4
CBM8	Binds to cellulose	336	83	0
CBM13	Binds to galactose residues, glycoside hydrolases and glycosyltransferases, xylan and GalNAc	235	533	87
CBM56	Binds to β-1,3-glucan	221	161	24
CBM23	Binds to mannan	212	66	8
CBM6	Bnds to cellulose, β-1,3-glucan, β-1,3-1,4-glucan and β-1,4-glucan.	209	276	63
CBM11	Binds to β-1,4-glucan and β-1,3-1,4-mixed linked glucans	155	93	0
CBM38	Binds to inulin	113	86	5

CBM61	Attached to various enzymes from families GH16, GH30, GH31, GH43, GH53 and GH66. Binds to β-1,4-galactan	111	61	22
CBM47	Binds to fucose	106	254	6
CBM5	Binds to chitin	70	71	9
CBM12	Binds to chitin	70	55	8
CBM26	Binds to starch	62	21	0
CBM30	Binds to cellulose	50	85	18
CBM62	Binds to galactose moieties found on xyloglucan, arabinogalactan and galactomannan	47	14	0
CBM59	Binds to mannan, xylan, and cellulose	46	17	0
CBM45	Binds to starch	29	38	0
CBM37	Broad binding specificity to xylan, chitin, microcrystalline and phosphoric-acid swollen cellulose, as well as more heterogeneous substrates, such as alfalfa cell walls, banana stem and wheat straw.	24	65	0
CBM70	Binds to hyaluronan	22	79	8
CBM77	Binds to pectin	15	16	0
CBM2	Binds to cellulose, chitin or xylan.	13	161	4
CBM42	Attached to various enzymes from GH54 family. Binds to arabinofuranose (present in arabinoxylan)	11	34	0
CBM54	Binds to xylan, yeast cell wall glucan and chitin	10	80	0
CBM63	Binds to cellulose	7	35	0

Carbohydrate Esterases (CEs)					
CAZy Family	Known CAZyme Activities	Known CAZyme Classification	HA + G	HA	G1
CE14	N-acetyl-1-D-myo-inosityl-2-amino-2-deoxy-α-D-glucopyranoside deacetylase; diacetylchitobiose deacetylase; mycothiol S-conjugate amidase	Debranching enzymes and Oligosaccharide-degrading enzymes	2683	3497	1108
CE12	pectin acetylesterase; rhamnogalacturonan acetylesterase; acetyl xylan esterase	Pectinases	1686	446	6
CE8	pectin methylesterase	Pectinases	1596	1120	469
CE6	acetyl xylan esterase	Debranching enzymes and Oligosaccharide-degrading enzymes	1123	362	68
CE15	4-O-methyl-glucuronoyl methylesterase	CAZy families processing fungal cell wall	791	666	38
CE2	acetyl xylan esterase	Debranching enzymes and Oligosaccharide-degrading enzymes	332	128	41

Polysaccharide Lyases (PLs)				
CAZy Family	Known CAZyme Activities	HA + G	HA	G1
PL9	pectate lyase; exopolygalacturonate lyase; thiopeptidoglycan lyase	1729	2083	190
PL12	heparin-sulfate lyase	1710	1124	116
PL8	hyaluronate lyase; chondroitin AC lyase; xanthan lyase; chondroitin ABC lyase	536	157	0
PL10	pectate lyase	363	176	10
PL11	rhamnogalacturonan lyase; exo-unsaturated rhamnogalacturonan lyase	304	71	0
PL17	alginate lyase; oligoalginate lyase	290	85	8
PL21	heparin lyase; heparin-sulfate lyase; acharan-sulfate lyase	40	4	0
PL7	poly(β-mannuronate) lyase / M-specific alginate lyase; α-L-guluronate lyase / G-specific alginate lyase; poly-(MG)-lyase / MG-specific alginate lyase	34	7	0
PL24	ulvan lyase	34	8	0
Other Modules				
CAZy Family		HA + G	HA	G1
cohesin		653	1005	81
dockerin		65	35	0
SLH		11	309	0

Table S5. 72 CAZy families with abundance values at least two times higher in the metagenomes containing humin (HM) versus those with only glucose (G).

Glycoside Hydrolases (GHs)					
CAZy Family	Known CAZyme Activities	Known CAZyme Classification	HM + G	HM	G2
GH43	β-xylosidase; α-L-arabinofuranosidase; arabinanase; xylanase; galactan 1,3-β-galactosidase; α-1,2-L-arabinofuranosidase; exo-α-1,5-L-arabinofuranosidase; [inverting] exo-α-1,5-L-arabinanase; β-1,3-xylosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	2181	1670	186
GH2	β-galactosidase; β-mannosidase; β-glucuronidase; α-L-arabinofuranosidase; mannosylglycoprotein endo-β-mannosidase; exo-β-glucosaminidase	Debranching enzymes and Oligosaccharide-degrading enzymes	2040	1567	193
GH5	endo-β-1,4-glucanase / cellulase; endo-β-1,4-xylanase; β-glucosidase; β-mannosidase; β-glucosylceramidase; glucan β-1,3-glucosidase; licheninase; exo-β-1,4-glucanase / cellodextrinase; glucan endo-1,6-β-glucosidase; mannan endo-β-1,4-mannosidase; cellulose β-1,4-cellobiosidase; steryl β-glucosidase; endoglycoceramidase; chitosanase; β-primeverosidase; xyloglucan-specific endo-β-1,4-glucanase; endo-β-1,6-galactanase; hesperidin 6-O-α-L-rhamnosyl-β-glucosidase; β-1,3-mannanase; arabinoxylan-specific endo-β-1,4-xylanase; mannan transglycosylase	Cellulases and endohemicellulases	1771	1684	579
GH28	polygalacturonase; exo-polygalacturonase; exo-polygalacturonosidase; rhamnogalacturonase; rhamnogalacturonan α-1,2-galacturonohydrolase; rhamnogalacturonan α-L-rhamnopyranohydrolase; endo-xylogalacturonan hydrolase	Pectinases	832	626	178
GH127	β-L-arabinofuranosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	660	305	24
GH105	unsaturated rhamnogalacturonyl hydrolase; d-4,5-unsaturated β-glucuronyl hydrolase	Pectinases	600	322	34

GH78	α-L-rhamnosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	571	599	80
GH92	mannosyl-oligosaccharide α-1,2-mannosidase; mannosyl-oligosaccharide α-1,3-mannosidase; mannosyl-oligosaccharide α-1,6-mannosidase; α-mannosidase; α-1,2-mannosidase; α-1,3-mannosidase; α-1,4-mannosidase; mannosyl-1-phosphodiester α-1,P-mannosidase	CAZy families processing fungal cell wall	501	228	49
GH20	β-hexosaminidase; lacto-N-biosidase; β-1,6-N-acetylglucosaminidase); β-6-SO3-N-acetylglucosaminidase	CAZy families processing fungal cell wall	499	245	31
GH9	endoglucanase; endo-β-1,3(4)-glucanase / lichenase-laminarinase; β-glucosidase; lichenase / endo-β-1,3-1,4-glucanase; exo-β-1,4-glucanase / cellodextrinase; cellobiohydrolase; xyloglucan-specific endo-β-1,4-glucanase / endo-xyloglucanase; exo-β-glucosaminidase	Cellulases and endohemicellulases	413	349	66
GH37	α,α-trehalase	Disaccharide Degrading	412	71	33
GH27	α-galactosidase; α-N-acetylgalactosaminidase; isomalto-dextranase; β-L-arabinopyranosidase; galactan:galactan galactosyltransferase	Debranching enzymes and Oligosaccharide-degrading enzymes	394	437	8
GH95	α-L-fucosidase; α-1,2-L-fucosidase; α-L-galactosidase	Cellulases and endohemicellulases	366	228	16
GH88	d-4,5-unsaturated β-glucuronyl hydrolase	Debranching enzymes and Oligosaccharide-degrading enzymes	364	142	39
GH8	chitosanase; cellulase; licheninase; endo-1,4-β-xylanase; reducing-end-xylose releasing exo-oligoxylanase	Cellulases and endohemicellulases	349	170	45
GH51	endoglucanase; endo-β-1,4-xylanase; β-xylosidase; α-L-arabinofuranosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	327	544	61
GH42	β-galactosidase; α-L-arabinopyranosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	318	200	78
GH97	glucoamylase; α-glucosidase; α-galactosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	316	177	18
GH29	α-L-fucosidase; α-1,3/1,4-L-fucosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	296	372	37
GH35	β-galactosidase; exo-β-glucosaminidase; exo-β-1,4-galactanase; β-1,3-galactosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	285	320	29

GH53	endo-β-1,4-galactanase	Debranching enzymes and Oligosaccharide-degrading enzymes	276	197	93
GH65	α,α-trehalase; maltose phosphorylase; trehalose phosphorylase; kojibiose phosphorylase; trehalose-6-phosphate phosphorylase; nigerose phosphorylase; 3-O-α-glucopyranosyl-L-rhamnose phosphorylase; 2-O-α-glucopyranosylglycerol: phosphate β-glucosyltransferase; α-glucosyl-1,2-β-galactosyl-L-hydroxylysine α-glucosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	269	151	6
GH26	β-mannanase; exo-β-1,4-mannobiohydrolase; β-1,3-xylanase; lichenase / endo-β-1,3-1,4-glucanase; mannobiose-producing exo-β-mannanase	Cellulases and endohemicellulases	266	343	0
GH55	exo-β-1,3-glucanase; endo-β-1,3-glucanase	CAZy families processing fungal cell wall	203	263	15
GH39	α-L-iduronidase; β-xylosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	190	360	67
GH10	endo-1,4-β-xylanase; endo-1,3-β-xylanase; tomatinase; xylan endotransglycosylase	Cellulases and endohemicellulases	189	286	33
GH99	glycoprotein endo-α-1,2-mannosidase; mannan endo-1,2-α-mannanase	CAZy families processing fungal cell wall	179	14	0
GH106	α-L-rhamnosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	166	126	0
GH76	α-1,6-mannanase	CAZy families processing fungal cell wall	161	56	6
GH115	xylan α-1,2-glucuronidase; α-(4-O-methyl)-glucuronidase	Debranching enzymes and Oligosaccharide-degrading enzymes	146	322	7
GH87	mycodextranase; α-1,3-glucanase	CAZy families processing fungal cell wall	130	257	0
GH54	α-L-arabinofuranosidase; β-xylosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	89	48	0
GH100	alkaline and neutral invertase	Disaccharide Degrading	76	52	19
GH67	α-glucuronidase; xylan α-1,2-glucuronidase	Debranching enzymes and Oligosaccharide-degrading enzymes	67	117	14
GH89	α-N-acetylglucosaminidase	CAZy families processing fungal cell wall	67	77	0
GH117	α-1,3-L-neoagarooligosaccharide hydrolase; α-1,3-L-neoagarobiase/neoagarobiose hydrolase	NA	66	30	7

GH125	exo-α-1,6-mannosidase	CAZy families processing fungal cell wall	64	29	8
GH110	α-galactosidase; α-1,3-galactosidase	NA	39	8	0
GH79	β-glucuronidase; hyaluronoglucuronidase; heparanase; baicalin β-glucuronidase; β-4-O-methyl-glucuronidase	Pectinases	34	8	0
GH123	β-N-acetylgalactosaminidase; glycosphingolipid β-N-acetylgalactosaminidase	NA	27	7	0
GH66	cycloisomaltooligosaccharide glucanotransferase; dextranase	NA	24	8	0
GH62	α-L-arabinofuranosidase	Debranching enzymes and Oligosaccharide-degrading enzymes	8	8	0
Glycosyl Transferases (GTs)					
CAZy Family	Known CAZyme Activities		HM + G	HM	G2
GT81	NDP-Glc: glucosyl-3-phosphoglycerate synthase; NDP-Man: mannosyl-3-phosphoglycerate synthase; ADP-Glc: glucosyl-2-glycerate synthase		591	204	99
GT94	GDP-Man: GlcA-β-1,2-Man-α-1,3-Glc-β-1,4-Glc-α-1-PP-undecaprenol β-1,4-mannosyltransferase		249	191	91
GT66	dolichyl-diphosphooligosaccharide—protein glycotransferase; undecaprenyl-diphosphooligosaccharide—protein glycotransferase		205	69	25
GT89	β-D-arabinofuranosyl-1-monophosphoryldecaprenol : arabinan β-1,2-arabinofuranosyltransferase		47	51	0
GT74	α-1,2-L-fucosyltransferase		17	70	0
GT92	UDP-Gal: N-glycan core α-1,6-fucoside β-1,4-galactosyltransferase; UDP-Gal: β-galactoside β-1,4-galactosyltransferase		16	73	0
GT22	Dol-P-Man: Man6GlcNAc2-PP-Dol α-1,2-mannosyltransferase; Dol-P-Man: Man8GlcNAc2-PP-Dol α-1,2-mannosyltransferase; Dol-P-Man: Man2-GlcNAc-phosphatidylinositol α-1,2-mannosyltransferase; Dol-P-Man: Man3-GlcNAc-phosphatidylinositol α-1,2-mannosyltransferase		4	7	0
Carbon Binding Modules (CBMs)					
CAZy Family	Known CAZyme Activities		HM + G	HM	G2
CBM32	Binds to galactose, lactose, polygalacturonic acid and LacNAc (β-D-galactosyl-1,4-β-D-N-acetylglucosamine)		1502	643	229
CBM6	Bnds to cellulose, β-1,3-glucan, β-1,3-1,4-glucan and β-1,4-glucan.		858	269	4
CBM67	Binds to L-rhamnose		427	472	56
CBM35	Binds to xylan and β-galactan, and decorated soluble mannans and mannooligosaccharides		416	117	13
CBM16	Binds to cellulose and glucomannan		358	196	80
CBM12	Binds to chitin		284	83	33
CBM57	binds to various glycosidases		227	167	0

CBM22	Binds to xylan and has affinity with mixed β-1,3/β-1,4-glucans		189	13	0
CBM51	Attached to various enzymes from families GH2, GH27, GH31, GH95, GH98 and GH101. Binds to galactose and to blood group A/B-antigens		160	80	22
CBM56	Binds to β-1,3-glucan		151	188	39
CBM4	Binds to xylan, β-1,3-glucan, β-1,3-1,4-glucan, β-1,6-glucan and amorphous cellulose but not with crystalline cellulose.		148	306	5
CBM40	Attached to various enzymes from GH 33 family. Binds to sialic acid		68	130	28
CBM30	Binds to cellulose		32	68	0
CBM53	Binds to starch		16	71	0
CBM54	Binds to xylan, yeast cell wall glucan and chitin		11	4	2
CBM38	Binds to inulin		9	12	0
Carbohydrate Esterases (CEs)					
CAZy Family	Known CAZyme Activities	Known CAZyme Classification	HM + G	HM	G2
CE2	acetyl xylan esterase	Debranching enzymes and Oligosaccharide-degrading enzymes	248	235	61
CE12	pectin acetylesterase; rhamnogalacturonan acetylesterase; acetyl xylan esterase	Pectinases	218	281	25
CE15	4-O-methyl-glucuronoyl methylesterase	CAZy families processing fungal cell wall	106	228	15
Polysaccharide Lyases (PLs)					
CAZy Family	Known CAZyme Activities		HM + G	HM	G2
PL9	pectate lyase; exopolygalacturonate lyase; thiopeptidoglycan lyase		292	223	77
PL11	rhamnogalacturonan lyase; exo-unsaturated rhamnogalacturonan lyase		71	9	0
PL10	pectate lyase		63	67	0
PL8	hyaluronate lyase; chondroitin AC lyase; xanthan lyase; chondroitin ABC lyase		47	106	0

Table S6. 13 aromatic degradation genes had abundance values at least two times higher in the metagenomes containing humic acid (HA) versus those with only glucose (G).

KEGG belonging to the Aromatics Degradation (AD) category			
KEGG Module	HA + G	HA	G1
Benzoyl-CoA degradation, benzoyl-CoA => 3-hydroxypimeloyl-CoA [PATH:map01220 map00362]	166	563	14
Cumate degradation, p-cumate => 2-oxopent-4-enoate + 2-methylpropanoate [PATH:map01220 map00622]	38	96	0
KEGG KO	HA + G	HA	G1
KO:K13954===yiaY===alcohol dehydrogenase [EC:1.1.1.1]	1430	1313	644
KO:K05710===hcaC===3-phenylpropionate/trans-cinnamate dioxygenase ferredoxin subunit	858	1107	313
KO:K00483===hpaB===4-hydroxyphenylacetate 3-monooxygenase [EC:1.14.14.9]	529	1920	16
KO:K07535===badH===2-hydroxycyclohexanecarboxyl-CoA dehydrogenase [EC:1.1.1.-]	265	237	87
KO:K00055===E1.1.1.90===aryl-alcohol dehydrogenase [EC:1.1.1.90]	201	470	5
KO:K00002===AKR1A1, adh===alcohol dehydrogenase (NADP+) [EC:1.1.1.2]	70	80	11
KO:K04115===bcrD, badG===benzoyl-CoA reductase subunit D [EC:1.3.7.8]	47	121	0
KO:K04112===bcrC, badD===benzoyl-CoA reductase subunit C [EC:1.3.7.8]	41	45	3
KO:K16050===hsaD===4,5:9,10-diseco-3-hydroxy-5,9,17-trioxoandrosta-1(10),2-diene-4-oate hydrolase [EC:3.7.1.17]	29	63	5
KO:K07539===oah===6-oxocyclohex-1-ene-carbonyl-CoA hydrolase [EC:3.7.1.21]	25	83	5
KO:K04114===bcrA, badF===benzoyl-CoA reductase subunit A [EC:1.3.7.8]	15	97	0
KO:K01612===E4.1.1.61===4-hydroxybenzoate decarboxylase [EC:4.1.1.61]	11	65	0
KO:K04113===bcrB, badE===benzoyl-CoA reductase subunit B [EC:1.3.7.8]	11	48	2

Table S7. Six aromatic degradation genes had abundance values at least two times higher in the metagenomes containing humin (HM) versus those with only glucose (G).

KEGG belonging to the Aromatics Degradation (AD) category			
KEGG Module	HM + G	HM	G2
Benzoyl-CoA degradation, benzoyl-CoA => 3-hydroxypimeloyl-CoA [PATH:map01220 map00362]	14	39	0
KEGG KO	HM + G	HM	G2
KO:K13953===adhP===alcohol dehydrogenase, propanol-preferring [EC:1.1.1.1]	1679	1297	500
KO:K14748===etbAa===ethylbenzene dioxygenase alpha subunit [EC:1.14.12.-]	56	30	0
KO:K10676===tfdB===2,4-dichlorophenol 6-monooxygenase [EC:1.14.13.20]	27	47	1
KO:K16049===hsaC===3,4-dihydroxy-9,10-secoandrosta-1,3,5(10)-triene-9,17-dione 4,5-dioxygenase [EC:1.13.11.25]	19	14	4
KO:K07534===badK===cyclohex-1-ene-1-carboxyl-CoA hydratase [EC:4.2.1.-]	15	17	0
KO:K14749===etbAb===ethylbenzene dioxygenase beta subunit [EC:1.14.12.-]	13	7	0

Table S8. Screening of *Geobacter* published and draft genomes from this study for aromatic degradation gene pathways.

KO Module ID	KO Module Name	KO Term ID	Geobacter metallireducens GS 15	Geobacter sulfurreducens AM 1	Geobacter sulfurreducens KN400	HA34 Bin15 Geobacter	HM12 Bin4 Geobacter	HM56 Bin7 Geobacter sulfurreducens
M00548	Benzene degradation benzene => catechol [PATH:map01220 map00362]	KO:K16249	0	0	0	0	0	0
		KO:K16250	0	0	0	0	0	0
		KO:K16251	0	0	0	0	0	0
		KO:K16252	0	0	0	0	0	0
		KO:K16253	0	0	0	0	0	0
		KO:K16254	0	0	0	2	0	0
M00418	Toluene degradation, anaerobic toluene => benzoyl-CoA [PATH:map01220 map00623]	KO:K16255	3	0	0	0	0	0
		KO:K16256	1	0	0	0	0	0
		KO:K16257	1	0	0	0	0	0
		KO:K16258	1	0	0	0	0	0
		KO:K16259	1	0	0	0	0	0
		KO:K16260	1	0	0	0	0	0
		KO:K16261	1	0	0	0	0	0
		KO:K16262	1	0	0	0	0	0
		KO:K16263	1	0	0	0	0	0
M00541	Benzoyl-CoA degradation benzoyl-CoA => 3-hydroxypimeloyl-CoA [PATH:map01220 map00362]	KO:K16264	0	0	0	0	0	0
		KO:K16265	0	0	0	0	0	0
		KO:K16266	0	0	0	0	0	0
		KO:K16267	0	0	0	0	0	0
		KO:K16268	1	0	0	0	0	0
		KO:K16269	2	0	0	0	0	0
		KO:K16270	1	0	0	0	0	0
		KO:K16271	1	0	0	0	0	0
		KO:K16272	2	0	0	0	0	0

Table S9. Screening of *Geothrix* published and draft genomes from this study for aromatic degradation gene pathways.

KO Module ID	KO Module Name	KO Term ID	Geothrix fermentans DSM 14018	HA34 Bin7 Geothrix fermentans	HM34 Bin13 Geothrix fermentans	Holophaga foetida TMBS4 DSM 6591	HM12 Bin10 Holophaga foetida	HM12 Bin22 Holophaga foetida	HM34 Bin11 Holophaga foetida	HM56 Bin28 Holophaga foetida
M00538	Toluene degradation toluene => benzoate [PATH:map01220 map00623]	KO:K15760	0	0	0	0	0	0	0	0
		KO:K15761	0	0	0	0	0	0	0	0
		KO:K15762	0	0	0	0	0	0	0	0
		KO:K15763	0	0	0	0	0	0	0	0
		KO:K15764	0	0	0	0	0	0	0	0
		KO:K15765	0	0	0	0	0	0	0	0
		KO:K00055	0	0	0	1	0	0	0	0
		KO:K00141	0	0	0	0	0	0	0	0
M00537	Xylene degradation xylene => methylbenzo	KO:K15757	0	0	0	0	0	0	0	0
		KO:K15758	0	0	0	0	0	0	0	0
		KO:K00055	0	0	0	1	0	0	0	0
		KO:K00141	0	0	0	0	0	0	0	0
M00568	Catechol ortho-cleavage catechol => 3-oxoadipate	KO:K03381	0	0	0	0	0	0	0	0
		KO:K01856	0	0	0	0	0	0	0	0
		KO:K03464	0	0	0	0	0	0	0	0
		KO:K01055	2	2	2	0	1	3	1	1
		KO:K14727	0	0	0	0	0	0	0	0
M00569	Catechol meta-cleavage catechol => acetyl-CoA / 4-methylcatechol => propanoyl-CoA [PATH:map01220	KO:K00446	0	0	0	0	0	0	0	0
		KO:K07104	0	0	0	0	0	0	0	0
		KO:K10217	0	0	0	0	0	0	0	0
		KO:K01821	1	3	1	1	1	1	1	0
		KO:K01617	0	0	0	0	0	0	0	0
		KO:K10216	0	0	0	0	0	0	0	0
		KO:K18364	0	0	0	0	0	0	0	0
		KO:K02554	0	0	0	0	0	0	0	0
		KO:K18365	0	0	0	0	0	0	0	0
		KO:K01666	0	0	0	0	0	2	1	1
		KO:K18366	0	0	0	0	0	0	0	0
		KO:K04073	0	0	0	0	1	0	0	0

Table S10. Screening of *Ignavibacterium/Melioribacter* published and draft genomes from this study for aromatic degradation gene pathways.

KO Module ID	KO Module Name	KO Term ID	Ignavibacterium album Mat9 16 JCM 16511	HA12 Bin10 Ignavibacteriales	HA12 Bin4 Ignavibacteriales	Melioribacter roseus P3M	HA12 Bin5 Melioribacter roseus
M00569	Catechol meta-cleavage catechol => acetyl-CoA / 4-methylcatechol => propanoyl-CoA [PATH:map01220 map00362 map00622]	KO:K00446	0	0	0	0	0
		KO:K07104	1	0	0	0	0
		KO:K10217	0	0	0	0	0
		KO:K01821	0	0	0	0	0
		KO:K01617	0	0	0	0	0
		KO:K10216	0	0	0	0	0
		KO:K18364	0	0	0	0	0
		KO:K02554	0	0	0	0	0
		KO:K18365	0	0	0	0	0
		KO:K01666	0	0	0	0	0
		KO:K18366	0	0	0	0	0
		KO:K04073	0	0	0	0	0

Table S11. Screening of *Anaeromyxobacter* published and draft genomes from this study for aromatic degradation gene pathways.

KO Module ID	KO Module Name	KO Term ID	Anaeromyxobacter dehalogenans 2CP 1	Anaeromyxobacter dehalogenans 2CP C	Anaeromyxobacter sp Fw109 5	Anaeromyxobacter sp K	HA34 Bin5 s Anaeromyxobacter dehalogens	HA56 Bin37 s Anaeromyxobacter dehalogens
M00537	Xylene degradation xylene => methylbenzoate [PATH:map01220 map00622]	KO:K15757	0	0	0	0	0	0
		KO:K15758	0	0	0	0	0	0
		KO:K00055	0	0	0	0	0	0
		KO:K00141	0	0	0	0	0	1
M00538	Toluene degradation toluene => benzoate [PATH:map01220 map00623]	KO:K15760	0	0	0	0	0	0
		KO:K15761	0	0	0	0	0	0
		KO:K15762	0	0	0	0	0	0
		KO:K15763	0	0	0	0	0	0
		KO:K15764	0	0	0	0	0	0
		KO:K15765	0	0	0	0	0	0
		KO:K00055	0	0	0	0	0	0
		KO:K00141	0	0	0	0	0	1
M00568	Catechol ortho-cleavage catechol => 3-oxoadipate [PATH:map01220 map00362]	KO:K03381	0	0	0	0	0	0
		KO:K01856	0	0	0	0	0	1
		KO:K03464	0	0	0	0	0	0
		KO:K01055	0	0	0	0	0	1
		KO:K14727	0	0	0	0	0	0
M00569	Catechol meta-cleavage catechol => acetyl-CoA / 4-methylcatechol => propanoyl-CoA [PATH:map01220 map00362 map00622]	KO:K00446	0	0	0	0	0	0
		KO:K07104	0	0	0	0	0	0
		KO:K10217	0	0	0	0	0	0
		KO:K01821	1	1	1	1	0	0
		KO:K01617	0	0	0	0	0	0
		KO:K10216	0	0	0	0	0	0
		KO:K18364	0	0	0	0	0	0
		KO:K02554	0	0	0	0	0	0
		KO:K18365	0	0	0	0	0	0
		KO:K01666	0	0	0	0	0	1
		KO:K18366	0	0	0	0	0	0
		KO:K04073	0	0	0	0	0	0

Table S12. Screening of *Prolixibacter* published and draft genomes from this study for aromatic degradation gene pathways.

KO Module ID	KO Module Name	KO Term ID	Prolixibacter bellariivorans ATCC BAA 1284	Prolixibacter bellariivorans JCM 13498	HM12 Bin18 Prolixibacter bellariivorans
M00569	Catechol meta-cleavage catechol => acetyl-CoA / 4-methylcatechol => propanoyl-CoA [PATH:map01220 map00362 map00622]	KO:K00446	0	0	0
		KO:K07104	0	0	0
		KO:K10217	0	0	0
		KO:K01821	0	0	0
		KO:K01617	0	0	0
		KO:K10216	0	0	0
		KO:K18364	0	0	0
		KO:K02554	0	0	0
		KO:K18365	0	0	0
		KO:K01666	0	0	1
		KO:K18366	0	0	0
		KO:K04073	0	0	0

References

(1) Richardson, D. J.; Butt, J. N.; Fredrickson, J. K.; Zachara, J. M.; Shi, L.; Edwards, M. J.; White, G.; Baiden, N.; Gates, A. J.; Marritt, S. J., The 'porin–cytochrome'model for microbe-to-mineral electron transfer. *Molecular microbiology* **2012,** *85*, (2), 201-212.

(2) Kyrpides, N., Geothrix fermentans DSM 14018. *Doe Joint Genome Institute. Retrieved October* **2012,** *26*.

(3) Coates, J.; Ellis, D.; Lovley, D., Geothrix fermentans gen. nov. sp. nov., an acetate-oxidizing Fe (III) reducer capable of growth via fermentation. *Int. J. Syst. Bacteriol* **1999,** *49*, 1615-1622.

(4) Nevin, K. P.; Lovley, D. R., Mechanisms for accessing insoluble Fe (III) oxide during dissimilatory Fe (III) reduction by Geothrix fermentans. *Applied and environmental microbiology* **2002,** *68*, (5), 2294-2299.

(5) Liu, Y.; Wang, Z.; Liu, J.; Levar, C.; Edwards, M. J.; Babauta, J. T.; Kennedy, D. W.; Shi, Z.; Beyenal, H.; Bond, D. R., A trans-outer membrane porin-cytochrome protein complex for extracellular electron transfer by *Geobacter sulfurreducens* PCA. *Environ. Microbiol. Rep.* **2014,** *6*, (6), 776-785.

(6) Cuskin, F.; Baslé, A.; Ladevèze, S.; Day, A. M.; Gilbert, H. J.; Davies, G. J.; Potocki-Véronèse, G.; Lowe, E. C., The GH130 family of mannoside phosphorylases contains glycoside hydrolases that target β-1, 2-mannosidic linkages in Candida mannan. *Journal of Biological Chemistry* **2015,** *290*, (41), 25023-25033.

(7) Blouzard, J.-C.; Bourgeois, C.; de Philip, P.; Valette, O.; Bélaïch, A.; Tardif, C.; Bélaïch, J.-P.; Pagès, S., Enzyme diversity of the cellulolytic system produced by Clostridium

cellulolyticum explored by two-dimensional analysis: identification of seven genes encoding new dockerin-containing proteins. *Journal of bacteriology* **2007,** *189*, (6), 2300-2309.

(8) Allgaier, M.; Reddy, A.; Park, J. I.; Ivanova, N.; D'haeseleer, P.; Lowry, S.; Sapra, R.; Hazen, T. C.; Simmons, B. A.; VanderGheynst, J. S., Targeted discovery of glycoside hydrolases from a switchgrass-adapted compost community. *PLoS One* **2010,** *5*, (1), e8812.

(9) Liu, N.; Zhang, L.; Zhou, H.; Zhang, M.; Yan, X.; Wang, Q.; Long, Y.; Xie, L.; Wang, S.; Huang, Y., Metagenomic insights into metabolic capacities of the gut microbiota in a fungus-cultivating termite (Odontotermes yunnanensis). *PLoS One* **2013,** *8*, (7), e69184.

(10) Munir, R. I.; Schellenberg, J.; Henrissat, B.; Verbeke, T. J.; Sparling, R.; Levin, D. B., Comparative analysis of carbohydrate active enzymes in Clostridium termitidis CT1112 reveals complex carbohydrate degradation ability. *PLoS One* **2014,** *9*, (8), e104260.

(11) Janeček, Š.; Svensson, B.; MacGregor, E. A., α-Amylase: an enzyme specificity found in various families of glycoside hydrolases. *Cellular and molecular life sciences* **2014,** *71*, (7), 1149-1170.

(12) Suyotha, W.; Yano, S.; Wakayama, M., α-1, 3-Glucanase: present situation and prospect of research. *World Journal of Microbiology and Biotechnology* **2016,** *32*, (2), 30.

(13) Suzuki, N.; Kim, Y.-M.; Fujimoto, Z.; Momma, M.; Okuyama, M.; Mori, H.; Funane, K.; Kimura, A., Structural elucidation of dextran degradation mechanism by Streptococcus mutans dextranase belonging to glycoside hydrolase family 66. *Journal of Biological Chemistry* **2012,** *287*, (24), 19916-19926.

(14) Geisler-Lee, J.; Geisler, M.; Coutinho, P. M.; Segerman, B.; Nishikubo, N.; Takahashi, J.; Aspeborg, H.; Djerbi, S.; Master, E.; Andersson-Gunnerås, S., Poplar carbohydrate-active enzymes. Gene identification and expression analyses. *Plant Physiology* **2006,** *140*, (3), 946-962.

(15) van Munster, J. M.; Daly, P.; Delmas, S.; Pullan, S. T.; Blythe, M. J.; Malla, S.; Kokolski, M.; Noltorp, E. C.; Wennberg, K.; Fetherston, R., The role of carbon starvation in the induction of enzymes that degrade plant-derived carbohydrates in Aspergillus niger. *Fungal genetics and biology* **2014,** *72*, 34-47.

(16) Mhuantong, W.; Charoensawan, V.; Kanokratana, P.; Tangphatsornruang, S.; Champreda, V., Comparative analysis of sugarcane bagasse metagenome reveals unique and conserved biomass-degrading enzymes among lignocellulolytic microbial communities. *Biotechnology for biofuels* **2015,** *8*, (1), 16.

(17) Zhao, Z.; Liu, H.; Wang, C.; Xu, J.-R., Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC genomics* **2013,** *14*, (1), 274.

(18) Chang, H.-X.; Yendrek, C. R.; Caetano-Anolles, G.; Hartman, G. L., Genomic characterization of plant cell wall degrading enzymes and in silico analysis of xylanses and polygalacturonases of Fusarium virguliforme. *BMC microbiology* **2016,** *16*, (1), 147.

(19) Wu, Y.-W.; Tang, Y.-H.; Tringe, S. G.; Simmons, B. A.; Singer, S. W., MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome* **2014,** *2*, (1), 26.

(20) Parks, D. H.; Imelfort, M.; Skennerton, C. T.; Hugenholtz, P.; Tyson, G. W., CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome research* **2015,** *25*, (7), 1043-1055.