

# Unearthing the microbial ecology of soil carbon cycling with DNA-SIP

Charles Pepe-Ranney and Ashley N Campbell \* †, Chantal Koechli †, Sean Berthrong ‡, and Daniel H Buckley † §

\*These authors contributed equally to the manuscript and should be considered co-first authors, †School of Integrative Plant Sciences, Cornell University, New York, USA, ‡Department of Biological Sciences, Butler University, and §corresponding author

## Abstract

We explored the dynamics of microbial contributions to decomposition in soil by coupling DNA Stable Isotope Probing (SIP) and high throughput DNA sequencing. Our experiment evaluated the degradative succession hypothesis, described dynamics of carbon (C) metabolism during organic matter degradation, and characterized bacteria that metabolize labile and structural C in soils. We added a complex amendment representing plant derived organic matter to soil substituting  $^{13}\text{C}$ -xylose or  $^{13}\text{C}$ -cellulose for unlabeled equivalents in two experimental treatments which were monitored for 30 days. Xylose and cellulose are abundant components in plant biomass and represent labile and structural C pools, respectively. We characterized 5,940 SSU rRNA gene operational taxonomic units (OTUs) finding evidence for  $^{13}\text{C}$ -incorporation into DNA from  $^{13}\text{C}$ -xylose and  $^{13}\text{C}$ -cellulose in 49 and 63 OTUs, respectively. The types of microorganisms that appeared  $^{13}\text{C}$ -labeled in the  $^{13}\text{C}$ -xylose treatment changed over time being predominantly *Firmicutes* at day 1 followed by *Bacteroidetes* at day 3 and then *Actinobacteria* at day 7. These dynamics of  $^{13}\text{C}$ -labeling suggest labile C traveled through different trophic levels within the soil bacterial community. In contrast, the microorganisms that metabolized cellulose-C increased in relative abundance later (after 14 days) with the highest number of OTUs exhibiting evidence for  $^{13}\text{C}$ -assimilation after 14 days. Microorganisms that metabolized cellulose-C belonged to cosmopolitan soil lineages that remain uncharacterized including *Spartobacteria*, *Chloroflexi* and *Planctomycetes*. Using an approach that reveals the C assimilation dynamics of specific microbial lineages we describe the ecological properties of functionally defined microbial groups that contribute to decomposition in soil.

stable isotope probing | structure-function relationships | soil microbial ecology | 16S rRNA gene

Abbreviations: C, Carbon; OTU, Operational Taxonomic Unit; SOM, Soil Organic Matter; BD, Buoyancy Density; SIP, Stable Isotope Probing

## Significance

Soil microorganisms drive C flux through the terrestrial biosphere, and models that predict terrestrial C flux can benefit by accounting for microbial ecophysiology in soils. However, characterizing the ecophysiology of microbes that mediate C decomposition in soil has proven difficult due to their overwhelming diversity. We characterized microbial C metabolism in soil and show that different types of C have distinct decomposition dynamics governed by different microbial lineages. For example, we found that uncharacterized microbial taxa, which are cosmopolitan in soils, assimilated cellulose-C into DNA. These microbes may drive cellulose decomposition on a global scale. We identify microbial lineages engaging in labile and structural C decomposition and explore their ecological properties.

## Introduction

Soils worldwide contain 2,300 Pg of carbon (C) which accounts for nearly 80% of the C present in the terrestrial biosphere [1, 2]. C respiration by soil microorganisms produces annually tenfold more  $\text{CO}_2$  than fossil fuel emissions [3]. Despite the contribution of microorganisms to global C flux, many global C models ignore the diversity of microbial physiology [4–6] and we still know little about the ecophysiology of soil microorganisms. Such knowledge should assist the development and refinement of global C models [7–10].

Most plant C is comprised of cellulose (30–50%) followed by hemicellulose (20–40%), and lignin (15–

25%) [11]. Hemicellulose, being the most soluble, degrades in the early stages of decomposition. Xylans are often an abundant component of hemicellulose, and xylans include differing amounts of xylose, glucose, arabinose, galactose, mannose, and rhamnose [12]. Xylose is often the most abundant sugar in hemicellulose, comprising as much as 60–90% of xylan in some plants (e.g. hardwoods [13], wheat [14], and switchgrass [15]). Microorganisms that respire labile C in the form of sugars proliferate during the initial stages of decomposition [16, 17], and metabolize as much as 75% of sugar C during the first 5 days [18]. In contrast, cellulose decomposition proceeds more slowly with rates increasing for approximately 15 days while degradation continues for 30–90 days [18, 19]. It is hypothesized that different microbial guilds mediate the decomposition of different plant biomass components [19–22]. The degradative succession hypothesis posits that fast growing organisms proliferate in response to the labile fraction of plant biomass such as sugars [23, 24] followed by slow growing organisms targeting structural C such as cellulose [23]. Evidence to support the degradative succession hypothesis comes from observing soil respiration dynamics and characterizing microorganisms cultured at different stages of decomposition. The degree to which the succession hypothesis presents an accurate model of litter decomposition has been questioned [21, 25, 26] and it’s clear that we need new approaches to dissect microbial contributions to C transformations in soils.

Though microorganisms mediate 80–90% of the soil C-cycle [27, 28], and microbial community composition can account for significant variation in C mineralization [29], terrestrial C-cycle models rarely consider the community composition of soils [30, 31]. Rates of soil C transformations are measured without knowledge of the organisms that mediate these reactions [28] leaving the importance of community membership towards maintaining ecosystem function undefined [28, 32, 33]. Variation in microbial community composition can be linked effectively to rates of soil processes when diagnostic genes for specific functions are available (e.g. denitrification [34], nitrification [35–37], methanotrophy [38], and nitrogen fixation [39]). However, the lack of diagnostic genes for describing soil-C transformations has limited progress in characterizing the contributions of individual microorganisms to decomposition. Remarkably, we still lack basic information on the physiology and ecology of the majority of organisms that live in soils. For example, contributions to soil processes remain uncharacterized for cosmopolitan bacterial phyla in soil such as *Acidobacteria*, *Chloroflexi*, *Planctomycetes*, and *Verrucomicrobia*. These phyla combined can comprise 32% of soil microbial commu-

nities (based on surveys of the SSU rRNA genes in soil) [40, 41].

Characterizing the functions of microbial taxa has relied historically on culturing microorganisms and subsequently characterizing their physiology in the laboratory, and on environmental surveys of genes diagnostic for specific processes. But, most microorganisms are difficult to grow in culture [40] and many biogeochemical processes lack suitable diagnostic genes. Nucleic acid stable-isotope probing (SIP) links genetic identity and activity without the need to grow microorganisms in culture and has expanded our knowledge of microbial contributions to biogeochemical processes [42]. However, nucleic acid SIP has notable complications including the need to add large amounts of labeled substrate [43], label dilution resulting in partial labeling of nucleic acids [43–45], the potential for cross-feeding and secondary label incorporation [45–50], and variation in genome G+C content [51–54]. As a result, most applications of SIP have targeted specialized microorganisms such as methanotrophs [43], methanogens [55], syntrophs [56], or microorganisms that target pollutants [57]. Exploring the soil-C cycle with SIP has proven to be more challenging because SIP has lacked the resolution necessary to characterize the specific contributions of individual microbial groups to the decomposition of plant biomass. High throughput DNA sequencing technology, however, improves the resolving power of SIP [58].

Coupling SIP with high throughput DNA sequencing now enables exploration of microbial C-cycling in soils. SSU rRNA amplicons are readily sequenced from numerous density gradient fractions across multiple samples thereby increasing the resolution of a typical nucleic acid SIP experiment [59]. It is now possible to use far less isotopically labeled substrate resulting in more environmentally realistic experimental conditions [58]. We have employed such a high resolution DNA stable isotope probing approach to explore the assimilation of xylose and/or cellulose into bacterial DNA in an agricultural soil.

Specifically, we added to soil a complex amendment representative of organic matter derived from fresh plant biomass. All treatments received the same amendment but the identity of isotopically labeled substrates was varied between treatments. Specifically, we set up a control treatment where all components were unlabeled, a treatment with <sup>13</sup>C-xylose instead of unlabeled xylose, and a treatment with <sup>13</sup>C-cellulose instead of unlabeled cellulose. Soil was sampled at days 1, 3, 7, 14, and 30 and we identified microorganisms that assimilated <sup>13</sup>C into DNA at each point in time. The experiment was designed to provide a test of the degradative succession hypothesis as it applies to soil bacteria,

to identify soil bacteria that metabolize xylose and cellulose, and to characterize temporal dynamics of xylose and cellulose metabolism in soil.

## Results

After adding the organic matter amendment to soil, we tracked the flow of  $^{13}\text{C}$  from  $^{13}\text{C}$ -xylose or  $^{13}\text{C}$ -cellulose into microbial DNA over time using DNA-SIP (Figure S1). The amendment consisted of compounds representative of plant biomass including cellulose, lignin, sugars found in hemicellulose, amino acids, and inorganic nutrients (see Supplemental Information (SI)). The amendment was added at 2.9 mg C g $^{-1}$  soil dry weight (d.w.), and this comprised 19% of the total C in the soil. The cellulose-C (0.88 mg C g $^{-1}$  soil d.w.) and xylose-C (0.42 mg C g $^{-1}$  soil d.w.) in the amendment comprised 6% and 3% of the total C in the soil, respectively. The soil microbial community respired 65% of the xylose within one day and 29% of the added xylose remained in the soil at day 30 (Figure S2). In contrast, cellulose-C declined at a rate of approximately 18  $\mu\text{g C d}^{-1} \text{g}^{-1}$  soil d.w. and 40% of added cellulose-C remained in the soil at day 30 (Figure S2).

**Community-level signal of  $^{13}\text{C}$ -assimilation in relation to substrate and time.** We assessed assimilation of  $^{13}\text{C}$  into microbial DNA by comparing the SSU rRNA gene sequence composition of SIP density gradient fractions between  $^{13}\text{C}$  treatments and the unlabeled control (see Methods and SI). In the control treatment, fraction density represented the majority of the variance in SSU rRNA gene composition (Figure 1). Genome G+C content correlates positively with DNA buoyant density and influences SSU rRNA gene composition in gradient fractions [51]. For the  $^{13}\text{C}$ -cellulose treatment, the SSU rRNA gene composition in gradient fractions deviated from control in high density fractions ( $> 1.72 \text{ g mL}^{-1}$ ) on days 14 and 30 (Figure 1). For the  $^{13}\text{C}$ -xylose treatment, SSU rRNA gene composition in gradient fractions also deviated from control in high density fractions, but it deviated from control on days 1, 3, and 7 (Figure 1). The SSU rRNA gene composition from the  $^{13}\text{C}$ -cellulose treatment and  $^{13}\text{C}$ -xylose treatment differed in high density gradient fractions indicating that different microorganisms were labeled in these treatments (Figure 1). Further, in the  $^{13}\text{C}$ -cellulose treatment, the SSU rRNA gene sequence composition in high density fractions was similar on days 14 and 30 indicating similar microorganisms had  $^{13}\text{C}$ -labeled DNA in  $^{13}\text{C}$ -cellulose treatments at days 14 and 30. In contrast, in the  $^{13}\text{C}$ -xylose treatment, the SSU rRNA gene composition of high density fractions varied between

days 1, 3, and 7 indicating that different microorganisms had  $^{13}\text{C}$ -labeled DNA on each of these days. In the  $^{13}\text{C}$ -xylose treatment, the SSU gene composition of high density fractions was similar to control on days 14 and 30 (Figure 1) indicating that  $^{13}\text{C}$  was no longer detectable in bacterial DNA on these days for this treatment.

### Temporal dynamics of OTU relative abundance in soil.

We monitored the soil microbial community over the course of the experiment by surveying SSU rRNA genes in non-fractionated DNA from the soil. The SSU rRNA gene composition of the non-fractionated DNA changed with time (Figure S3, P-value = 0.023,  $R^2 = 0.63$ , Adonis test [60]). In contrast, the microbial community could not be shown to change with treatment (P-value 0.23, Adonis test) (Figure S3). The latter result demonstrates the substitution of  $^{13}\text{C}$ -labeled substrates for unlabeled equivalents could not be shown to alter the soil microbial community composition. Twenty-nine OTUs exhibited sufficient statistical evidence (adjusted P-value < 0.10, Wald test) to conclude they changed in relative abundance in the non-fractionated DNA over the course of the experiment (Figure S4). When SSU rRNA gene abundances were combined at the taxonomic rank of “class”, the classes that changed in abundance (adjusted P-value < 0.10, Wald test) were the *Bacilli* (decreased), *Flavobacteria* (decreased), *Gamma proteobacteria* (decreased), and *Herpetosiphonales* (increased) (Figure S5). Of the 29 OTUs that changed in relative abundance over time, 14 putatively incorporated  $^{13}\text{C}$  into DNA (Figure S4). OTUs that likely assimilated  $^{13}\text{C}$  from  $^{13}\text{C}$ -cellulose tended to increase in relative abundance with time whereas OTUs that assimilated  $^{13}\text{C}$  from  $^{13}\text{C}$ -xylose tended to decrease (Figure S6). OTUs that responded to both substrates did not exhibit a consistent relative abundance response over time as a group (Figure S4 and S6).

**Changes in the phylogenetic composition of  $^{13}\text{C}$ -labeled OTUs with substrate and time.** If an OTU exhibited strong evidence for assimilating  $^{13}\text{C}$  into DNA, we refer to that OTU as a “responder” (see Methods and SI for our operational definition of “responder”). The SSU rRNA gene sequences produced in this study were binned into 5,940 OTUs and we assessed evidence of  $^{13}\text{C}$ -labeling from both  $^{13}\text{C}$ -cellulose and  $^{13}\text{C}$ -xylose for each OTU. Forty-one OTUs responded to  $^{13}\text{C}$ -xylose, 55 OTUs responded to  $^{13}\text{C}$ -cellulose, and 8 OTUs responded to both xylose and cellulose (Figure 2, Figure 3, Figure S7, Table S1, and Table S2). The number of xylose responders peaked at days 1 and 3 and declined with time. In contrast, the number of cel-

<sup>310</sup> luloose responders increased with time peaking at days 14 and 30 (Figure S8).

The phylogenetic composition of xylose responders changed with time (Figure 2 and Figure 4) and <sup>315</sup> 86% of xylose responders shared > 97% SSU rRNA gene sequence identity with bacteria cultured in isolation (Table S1). On day 1, *Bacilli* OTUs represented 84% of xylose responders (Figure 4) and the majority of these OTUs were closely related to <sup>320</sup> cultured representatives of the genus *Paenibacillus* (Table S1, Figure 3). For example, “OTU.57” (<sup>325</sup> Table S1), annotated as *Paenibacillus*, had a strong signal of <sup>13</sup>C-labeling at day 1 coinciding with its maximum relative abundance in non-fractionated DNA. The relative abundance of “OTU.57” declined until day 14 and “OTU.57” did not appear to be <sup>13</sup>C-labeled after day 1 (Figure S9). On day 3, *Bacteroidetes* OTUs comprised 63% of xylose responders (Figure 4) and these OTUs were <sup>330</sup> closely related to cultured representatives of the *Flavobacteriales* and *Sphingobacteriales* (Table S1, Figure 3). For example, “OTU.14”, annotated as a flavobacterium, had a strong signal for <sup>13</sup>C-labeling in the <sup>13</sup>C-xylose treatment at days 1 and 3 coinciding with its maximum relative abundance <sup>335</sup> in non-fractionated DNA. The relative abundance of “OTU.14” then declined until day 14 and did not show evidence of <sup>13</sup>C-labeling beyond day 3 (Figure S9). Finally, on day 7, *Actinobacteria* OTUs represented 53% of the xylose responders <sup>340</sup> (Figure 4) and these OTUs were closely related to cultured representatives of *Micrococcales* (Table S1, Figure 3). For example, “OTU.4”, annotated as *Agromyces*, had signal for <sup>13</sup>C-labeling in the <sup>13</sup>C-xylose treatment on days 1, 3 and 7 with <sup>345</sup> the strongest evidence of <sup>13</sup>C-labeling at day 7 and did not appear <sup>13</sup>C-labeled at days 14 and 30. The relative abundance of “OTU.4” in non-fractionated DNA increased until day 3 and then declined until day 30 (Figure S9). *Proteobacteria* were also <sup>350</sup> common among xylose responders at day 7 where they comprised 40% of xylose responder OTUs. Notably, *Proteobacteria* represented the majority (6 of 8) of OTUs that responded to both cellulose and xylose (Figure S7).

<sup>355</sup> The phylogenetic composition of cellulose responders did not change with time to the same extent as the xylose responders. Also, in contrast to xylose responders, cellulose responders often were not closely related (< 97% SSU rRNA <sup>360</sup> gene sequence identity) to cultured isolates. Both the relative abundance and the number of cellulose responders increased over time peaking at days 14 and 30 (Figure 2, Figure S8, and Figure S6). Cellulose responders belonged to the *Pro-<sup>365</sup>* teobacteria (46%), *Verrucomicrobia* (16%), *Planctomycetes* (16%), *Chloroflexi* (8%), *Bacteroidetes*

(8%), *Actinobacteria* (3%), and *Melanobacteria* (1 OTU) (Table S2).

The majority (85%) of cellulose responders <sup>370</sup> outside of the *Proteobacteria* shared < 97% SSU rRNA gene sequence identity to bacteria cultured in isolation. For example, 70% of the *Verrucomicrobia* cellulose responders fell within unidentified *Spartobacteria* clades (Figure 3), and these shared <sup>375</sup> < 85% SSU rRNA gene sequence identity to any characterized isolate. The *Spartobacteria* OTU “OTU.2192” exemplified many cellulose responders (Table S2, Figure S9). “OTU.2192” increased in non-fractionated DNA relative abundance with <sup>380</sup> time and evidence for <sup>13</sup>C-labeling of “OTU.2192” in the <sup>13</sup>C-cellulose treatment increased over time with the strongest evidence at days 14 and 30 (Figure S9). Most *Chloroflexi* cellulose responders belonged to an unidentified clade within the *Herpetosiphonales* (Figure 3) and they shared < 89% SSU rRNA gene sequence identity to any characterized isolate. Characteristic of *Chloroflexi* cellulose responders, “OTU.64” increased in relative abundance over 30 days and evidence for <sup>13</sup>C-labeling of “OTU.64” in the <sup>13</sup>C-cellulose treatment peaked days 14 and 30 (Figure S9). *Bacteroidetes* cellulose responders fell within the *Cytophagales* in contrast with *Bacteroidetes* xylose responders that belonged instead to the *Flavobacteriales* or *Sphingobacteriales* (Figure 3). *Bacteroidetes* cellulose responders included one OTU that shared 100% SSU rRNA gene sequence identity to a *Sporocytophaga* species, a genus known to include cellulose degraders. The majority (86%) of cellulose responders in the *Proteobacteria* were <sup>390</sup> closely related (> 97% identity) to bacteria cultured in isolation, including representatives of the genera: *Cellvibrio*, *Devosia*, *Rhizobium*, and *Sorangium*, which are all known for their ability to degrade cellulose (Table S2). Proteobacterial cellulose responders belonged to *Alpha* (13 OTUs), *Beta* (4 OTUs), *Gamma* (5 OTUs), and *Delta-*<sup>395</sup> *proteobacteria* (6 OTUs).

#### Characteristics of cellulose and xylose responders.

Cellulose responders, relative to xylose responders, tended to have lower relative abundance in non-fractionated DNA, demonstrated signal consistent with higher atom % <sup>13</sup>C in labeled DNA, and had lower estimated *rrn* copy number (Figure 5). In the non-fractionated DNA, cellulose responders had lower relative abundance ( $1.2 \times 10^{-3}$  (s.d.  $3.8 \times 10^{-3}$ ) than xylose responders ( $3.5 \times 10^{-3}$  (s.d.  $5.2 \times 10^{-3}$ )) (Figure 4, P-value =  $1.12 \times 10^{-5}$ , Wilcoxon Rank Sum test). Six of the ten most common OTUs observed in the non-fractionated DNA responded to xylose, and, seven of the ten most abundant responders to xylose or cellulose in the non-fractionated DNA were xylose responders.

DNA buoyant density (BD) increases in proportion to atom %  $^{13}\text{C}$ . Hence, the extent of  $^{13}\text{C}$  incorporation into DNA can be evaluated by the difference in BD between  $^{13}\text{C}$ -labeled and unlabeled DNA. We calculated for each OTU its mean BD weighted by relative abundance to determine its “center of mass” within a given density gradient. We then quantified for each OTU the difference in center of mass between control gradients and gradients from  $^{13}\text{C}$ -xylose or  $^{13}\text{C}$ -cellulose treatments (see SI for the detailed calculation, Figure S11). We refer to the change in center of mass position for an OTU in response to  $^{13}\text{C}$ -labeling as  $\Delta\hat{B}\text{D}$ .  $\Delta\hat{B}\text{D}$  can be used to compare relative differences in  $^{13}\text{C}$ -labeling between OTUs.  $\Delta\hat{B}\text{D}$  values, however, are not comparable to the BD changes observed for DNA from pure cultures both because they are based on relative abundance in density gradient fractions (and not DNA concentration) and because isolated strains grown in uniform conditions generate uniformly labeled molecules while OTUs composed of heterogeneous strains in complex environmental samples do not. Cellulose responder  $\Delta\hat{B}\text{D}$  ( $0.0163 \text{ g mL}^{-1}$  (s.d. 0.0094)) was greater than that of xylose responders ( $0.0097 \text{ g mL}^{-1}$  (s.d. 0.0094)) (Figure 5, P-value =  $1.8610 \times 10^{-6}$ , Wilcoxon Rank Sum test).

We predicted the *rrn* gene copy number for responders as described [61]. The ability to proliferate after rapid nutrient influx correlates positively to a microorganism’s *rrn* copy number [62]. Cellulose responders possessed fewer estimated *rrn* copy numbers ( $2.7$  (1.2 s.d.)) than xylose responders ( $6.2$  (3.4 s.d.)) ( $P = 1.878 \times 10^{-9}$ , Wilcoxon Rank Sum test, Figure 5 and Figure S10). Furthermore, the estimated *rrn* gene copy number for xylose responders was inversely related to the day of first response ( $P = 2.02 \times 10^{-15}$ , Wilcoxon Rank Sum test, Figure S10, Figure 5).

We assessed phylogenetic clustering of  $^{13}\text{C}$ -responsive OTUs with the Nearest Taxon Index (NTI) and the Net Relatedness Index (NRI) [63]. We also quantified the average clade depth of cellulose and xylose responders with the consenTRAIT metric [64]. Briefly, the NRI and NTI evaluate phylogenetic clustering against a null model for the distribution of a trait in a phylogeny. The NRI and NTI values are z-scores or standard deviations from the mean and thus the greater the magnitude of the NRI/NTI, the stronger the evidence for clustering (positive values) or overdispersion (negative values). NRI assesses overall clustering whereas the NTI assesses terminal clustering [65]. The consenTRAIT metric is a measure of the average clade depth for a trait in a phylogenetic tree. NRI values indicate that cellulose responders clustered overall and at the tips of the phylogeny (NRI: 4.49, NTI: 1.43) while xylose responders

clustered terminally (NRI: -1.33, NTI: 2.69). The consenTRAIT clade depth for xylose and cellulose responders was 0.012 and 0.028 SSU rRNA gene sequence dissimilarity, respectively. As reference, the average clade depth is approximately 0.017 SSU rRNA gene sequence dissimilarity for arabinose (arabinose like xylose is a five C sugar found in hemicellulose) utilization as inferred from genomic analyses, and was 0.013 and 0.034 SSU rRNA gene sequence dissimilarity for glucosidase and cellulase genomic potential, respectively [64, 66]. These results indicate xylose responders form terminal clusters dispersed throughout the phylogeny while cellulose responders form deep clades of terminally clustered OTUs.

## Discussion

We identified microorganisms participating in soil C cycling using a nucleic acid SIP approach. Specifically, we observed assimilation of  $^{13}\text{C}$  from either  $^{13}\text{C}$ -xylose or  $^{13}\text{C}$ -cellulose into DNA for 104 OTUs in an agricultural soil. We found  $^{13}\text{C}$  from  $^{13}\text{C}$ -xylose appeared to move into and then out of groups of related OTUs over time. By coupling nucleic acid SIP to high throughput sequencing we could diagnose OTU activity even when OTUs were at low relative abundance in non-fractionated DNA (e.g. on three occasions we did not detect  $^{13}\text{C}$ -responders in the non-fractionated DNA). Our results support the degradative succession hypothesis, elucidate ecophysiological properties of soil microorganisms, reveal activity of widespread uncultured soil bacteria, and begin to piece together the microbial food web in soils.

The degradative succession hypothesis predicts an ecological transition in activity during the decomposition of labile and structural plant organic matter. Our results concur with the degradative succession hypothesis. Microorganisms rapidly metabolized xylose-C while cellulose-C metabolism proceeded slowly. Xylose is a major constituent of hemicellulose and is a labile component of fresh plant biomass. The phylogenetic composition of xylose responders changed between days 1, 3 and 7 and few OTUs appeared  $^{13}\text{C}$ -labeled in the  $^{13}\text{C}$ -xylose treatment after day 7. In the  $^{13}\text{C}$ -cellulose treatment,  $^{13}\text{C}$ -labeled OTUs were few in the beginning of the experiment and most abundant on days 14 and 30. Finally, few (8 of 104) OTUs appeared to metabolize both xylose and cellulose indicating most  $^{13}\text{C}$ -responders had distinct activity.

The ecological characteristics of microorganisms are often inferred from correlations between community composition and environmental characteristics [67]. In this experiment, we identified ecological groups as a function of *in situ* metabolism and

inferred the ecological properties of these groups through temporal dynamics of  $^{13}\text{C}$ -assimilation, 540 the extent of OTU  $^{13}\text{C}$ -labeling, and phylogenetic affiliation. Xylose responders grew faster than cellulose responders and appeared to assimilate C from multiple sources. Xylose responders assimilated xylose-C into DNA within 24 hours and 545 had low  $\Delta\bar{BD}$  relative to cellulose responders suggesting xylose was not the sole C source used for growth. Xylose represented 15% of the amendment and 3% of total soil C. Xylose responders often included the most abundant OTUs within 550 the non-fractionated DNA and had high estimated *rrn* copy number relative to cellulose responders. However, to some degree, high *rrn* gene copy number may inflate observed xylose responder relative abundance. Notably, the majority of xylose 555 responder SSU rRNA genes (86%) matched SSU rRNA genes from cultured isolates at high sequence identity (> 97%).

Cellulose responders, on the other hand, incorporated  $^{13}\text{C}$  into DNA after xylose responders and 560 appeared to specialize in using cellulose as a C source. Cellulose responders grew over a span of weeks and had high  $\Delta\bar{BD}$  indicating cellulose remained their dominant C source even though multiple C sources were present (cellulose represented 565 6% of total C present in soil at the start of the experiment). Cellulose responders were also lower in relative abundance on average within the non-fractionated DNA and had lower estimated *rrn* copy number than xylose responders. The majority 570 of cellulose responders were not close relatives of cultured isolates although a number of cellulose responders shared high SSU rRNA gene sequence identity with cultured *Proteobacteria* (e.g. *Celvibrio*). We identified cellulose responders among 575 phyla such as *Verrucomicrobia*, *Chloroflexi*, and *Planctomycetes* – common soil phyla whose functions within soil communities remain unknown.

*Verrucomicrobia* represented 16% of the cellulose responders. *Verrucomicrobia* are cosmopolitan 580 than soil microorganisms [68] that can make up to 23% of SSU rRNA gene sequences in soils [68] and 9.8% of soil SSU rRNA [69]. Genomic analyses and laboratory experiments show that various isolates within the *Verrucomicrobia* are capable of 585 methanotrophy, diazotrophy, and cellulose degradation [70, 71]. Moreover, *Verrucomicrobia* have been hypothesized to degrade polysaccharides in many environments [72–74]. However, only one of the 15 most abundant verrucomicrobial phylotypes 590 in globally distributed soil samples shared > 93% SSU rRNA gene sequence identity with a cultured isolate [68] and hence the role of soil *Verrucomicrobia* in global C-cycling remains unknown. The majority of verrucomicrobial cellulose 595 responders belonged to two clades that fall within

the *Spartobacteria* (Figure 3). *Spartobacteria* outnumbered all other *Verrucomicrobia* phylotypes in SSU rRNA gene surveys of 181 globally distributed soil samples [68]. Given their ubiquity and abundance in soil as well as their demonstrated incorporation of  $^{13}\text{C}$  from  $^{13}\text{C}$ -cellulose, *Verrucomicrobia* lineages, particularly *Spartobacteria*, may be important contributors to cellulose decomposition on a global scale.

Other notable cellulose responders include OTUs in the *Planctomycetes* and *Chloroflexi* both of which have previously been shown to assimilate  $^{13}\text{C}$  from  $^{13}\text{C}$ -cellulose added to soil [75]. *Planctomycetes* are common in soil [40], comprising 4 to 7% of bacterial cells in many soils [76, 77] and 7% ± 5% of SSU rRNA [78]. Although soil *Planctomycetes* are widespread, their activities in soil remain uncharacterized. *Planctomycetes* represented 16% of cellulose responders and shared < 92% 615 SSU rRNA gene sequence identity to their most closely related cultured isolates. *Chloroflexi* are known for metabolically diverse lifestyles ranging from anoxygenic phototrophy to organohalide respiration [79] and are among the six most abundant 620 bacterial phyla in soil [40]. Recent studies have focused on *Chloroflexi* roles in C cycling [79–81] and several *Chloroflexi* isolates use cellulose [79–81]. Four of the five *Chloroflexi* cellulose responders belong to a single clade within the *Herpetosiphonales* 625 (Figure 3).

Finally, a single cellulose responder belonged to the *Melanabacteria* phylum (95% shared SSU rRNA gene sequence identity with *Vampirovibrio chlorellavorus*). The phylogenetic position of 630 *Melanabacteria* is debated but *Melanabacteria* have been proposed to be a non-phototrophic sister phylum to *Cyanobacteria*. An analysis of a *Melanabacteria* genome [82] suggests the genomic capacity to degrade polysaccharides though *Vampirovibrio chlorellavorus* is an obligate predator of green 635 algae [83].

Responders did not necessarily assimilate  $^{13}\text{C}$  directly from  $^{13}\text{C}$ -xylose or  $^{13}\text{C}$ -cellulose but, in many ways, knowledge of secondary C degradation 640 and/or microbial biomass turnover may be more interesting with respect to the soil C-cycle than knowledge of primary degradation. The response to xylose suggests xylose-C moved through different trophic levels within the soil bacterial food web. The *Bacilli* degraded xylose first (65% of the xylose-C had been respired by day 1) representing 84% of day 1 xylose responders. *Bacilli* also comprised about 6% of SSU rRNA genes present in non-fractionated DNA on day 1. However, few *Bacilli* remained  $^{13}\text{C}$ -labeled by day 3 and their abundance declined reaching about 2% of soil SSU rRNA genes by day 30. Members of the 645 *Bacillus* [84] and *Paenibacillus* in particular [59]

have been previously implicated as labile C decomposers. The decline in relative abundance of *Bacilli* could be attributed to mortality and/or sporulation coupled to mother cell lysis. *Bacteroidetes* OTUs appeared  $^{13}\text{C}$ -labeled at day 3 concomitant with the decline in relative abundance and loss of  $^{13}\text{C}$ -label for *Bacilli*. Finally, *Actinobacteria* appeared  $^{13}\text{C}$ -labeled at day 7 as *Bacteroidetes* xylose responders declined in relative abundance and became unlabeled. Hence, it seems reasonable to propose that *Bacteroidetes* and *Actinobacteria* xylose responders became labeled via the consumption of  $^{13}\text{C}$  derived from  $^{13}\text{C}$ -labeled microbial biomass as opposed to primary degradation of  $^{13}\text{C}$ -xylose.

The inferred physiology of *Actinobacteria* and *Bacteroidetes* xylose responders provides further evidence for C transfer by saprotrophy and/or predation. Most of the *Actinobacteria* xylose responders that appeared  $^{13}\text{C}$ -labeled at day 7 were members of the *Micrococcales* (Figure 3) and the most abundant  $^{13}\text{C}$ -labeled *Micrococcales* OTU at day 7 (OTU.4, Table S1) is annotated as belonging in the *Agromyces*. *Agromyces* are facultative predators that feed on the gram-positive *Luteobacter* in culture [85]. Additionally, certain types of *Bacteroidetes* can assimilate  $^{13}\text{C}$  from  $^{13}\text{C}$ -labeled *Escherichia coli* added to soil [86]. Alternatively, it is possible that *Bacilli*, *Bacteroidetes*, and *Actinobacteria* are adapted to use xylose at different concentrations and that the observed activity dynamics resulted from changes in xylose concentration over time and/or that *Actinobacteria* and *Bacteroidetes* xylose responders consumed waste products generated by primary xylose metabolism (e.g. organic acids produced during xylose metabolism). These latter two hypotheses cannot explain the sequential loss of  $^{13}\text{C}$ -label in combination with the abundance dynamics in non-fractionated DNA, however. If trophic transfer caused the activity dynamics, at least three different ecological groups exchanged C in 7 days. Models of the soil C cycle often exclude trophic interactions between soil bacteria (e.g. [87]), yet when soil C models do account for predators and/or saprophytes, trophic interactions are predicted to have significant effects on the fate of soil C [88].

**Implications for soil C cycling models.** Functional niche characterization for soil microorganisms is necessary to predict whether and how biogeochemical processes vary with microbial community composition. Functional niches are defined by soil microbiologists and have been successfully incorporated into biogeochemical process models (E.g. [88, 89]). In some C models ecological strategies such as growth rate and substrate specificity are parameters for functional niche behavior [88].

The phylogenetic breadth of a functionally defined

group is often inferred from the distribution of diagnostic genes across genomes [66] or from the physiology of isolates cultured on laboratory media [64]. For instance, the wide distribution of the glycolysis operon in microbial genomes is interpreted as evidence that many soil microorganisms participate in glucose turnover [9]. However, the functional niche may depend less on the distribution of diagnostic genes across genomes and more on life history traits that allow organisms to compete for a given substrate as it occurs in the soil. For instance, fast growth and rapid resuscitation allow microorganisms to compete for labile C which may often be transient in soil. Hence, life history traits may constrain the diversity of microorganisms that metabolize a given C source in the soil under a given set of conditions.

Biogeochemical processes mediated by a broad array of taxa are assumed insensitive to community change relative to processes mediated by a narrow suite of microorganisms [9, 90]. In addition, the diversity of a functionally defined group engaged in a specific C transformation is expected to correlate positively with C lability [9]. However, the diversity of labile C and structural C decomposers in soil has not been quantified directly. We found comparable numbers of OTUs responded to  $^{13}\text{C}$ -cellulose and  $^{13}\text{C}$ -xylose (63 and 49, respectively). Cellulose responders were phylogenetically clustered suggesting that the ability to degrade cellulose is phylogenetically conserved. The clade depth of cellulose responders, 0.028 SSU rRNA gene sequence dissimilarity, is on the same order as that observed for glycoside hydrolases which are diagnostic enzymes for cellulose degradation [66]. Xylose responders clustered in terminal branches indicating groups of closely related taxa metabolized xylose but xylose responders also clustered phylogenetically with respect to time of response (Figure 3, Figure 4). For example, xylose responders on day 1 are dominated by members of *Paenibacillus*. Thus, microorganisms that degraded labile C and structural C were both limited in diversity. Although the genes for xylose metabolism are likely widespread in the soil community, it's possible only a limited diversity of organisms had the ecological characteristics required to degrade xylose under experimental conditions. Therefore it's possible that only a limited number of taxa actually participate in the metabolism of labile C-sources under a given set of conditions, and hence changes in community composition may alter the dynamics of structural and labile C-transformations in soil.

Broadly, we observed labile C use by fast growing generalists and structural C use by slow growing specialists. These results agree with the MIMICS model which simulates leaf litter decomposition by modeling microbial decomposers as two func-

tionally defined groups, copiotrophs or oligotrophs [89]. Including these functional types improved predictions of C storage in response to environmental change. We identified microbial lineages engaged in labile and structural C decomposition that can be defined as copiotrophs or oligotrophs, respectively. Additionally, we show that soil-C may travel through multiple bacterial trophic levels where each C transfer represents an opportunity for C stabilization in association with soil minerals or C loss by respiration. Our understanding of soil C dynamics will likely improve as we develop a more granular understanding of the ecological diversity of microorganisms that mediate C transformations in soil.

### Conclusion.

Microorganisms govern C-transformations in soil influencing climate change on a global scale but we do not know the identities of microorganisms that carry out specific transformations. In this experiment microorganisms from physiologically uncharacterized but cosmopolitan soil lineages participated in cellulose decomposition. Cellulose responders included members of the *Verrucomicrobia* (*Spartobacteria*), *Chloroflexi*, *Bacteroidetes* and *Planctomycetes*. *Spartobacteria* in particular are globally cosmopolitan soil microorganisms and are often the most abundant *Verrucomicrobia* order in soil [68]. Fast-growing aerobic spore formers from *Firmicutes* assimilated labile C in the form of xylose. Xylose responders within the *Bacteroidetes* and *Actinobacteria* likely became labeled by consuming <sup>13</sup>C-labeled constituents of microbial biomass either by saprotrophy or predation. Our results suggest that cosmopolitan *Spartobacteria* may degrade cellulose on a global scale, decomposition of labile plant C may initiate trophic transfer within the bacterial food web, and life history traits may act as a filter constraining the diversity of active microorganisms relative to those with the genomic potential for a given metabolism.

### Methods

All code to take raw SSU rRNA gene sequencing reads to final publication figures and through all presented analyses is located at the following URL: [https://github.com/chuckpr/CSIP\\_succession\\_data\\_analysis](https://github.com/chuckpr/CSIP_succession_data_analysis).

DNA sequences are deposited on MG-RAST (Accession XXXXXXXX).

Twelve soil cores (5 cm diameter x 10 cm depth) were collected from six sampling locations within an organically managed agricultural field in Penn Yan, New York. Soils were sieved (2 mm), homogenized, distributed into flasks (10 g in each 250 ml flask, n = 36) and equilibrated for 2 weeks. We amended soils with a mixture containing 2.9 mg C

<sup>825</sup> g<sup>-1</sup> soil dry weight (d.w.) and brought soil to 50% water holding capacity. By mass the amendment contained 38% cellulose, 23% lignin, 20% xylose, 3% arabinose, 1% galactose, 1% glucose, and 0.5% mannose. 10.6% amino acids (Teknova C9795) and 2.9% Murashige Skoog basal salt mixture which contains macro and micro-nutrients that are associated with plant biomass (Sigma Aldrich M5524). This mixture approximates the molecular composition of switchgrass biomass with hemicellulose replaced by its constituent monomers [91]. We set up three parallel treatments varying the isotopically labeled component in each treatment. The treatments were (1) a control treatment with all unlabeled components, (2) a treatment with <sup>13</sup>C-cellulose instead of unlabeled cellulose (synthesized as described in SI), and (3) a treatment with <sup>13</sup>C-xylose (98 atom% <sup>13</sup>C, Sigma Aldrich) instead of unlabeled xylose. Other details relating to substrate addition can be found in SI. Microcosms were sampled destructively at days 1 (control and xylose only), 3, 7, 14, and 30 and soils were stored at -80°C until nucleic acid extraction. The abbreviation 13CXPS refers to the <sup>13</sup>C-xylose treatment (<sup>13</sup>C Xylose Plant Simulant), 13CCPS refers to the <sup>13</sup>C-cellulose treatment, and 12CCPS refers to the control treatment.

We used DESeq2 (R package), an RNA-Seq differential expression statistical framework [92], to identify OTUs that were enriched in high density gradient fractions from <sup>13</sup>C-treatments relative to corresponding gradient fractions from control treatments (for review of RNA-Seq differential expression statistics applied to microbiome OTU count data see [93]). We define “high density gradient fractions” as gradient fractions whose density falls between 1.7125 and 1.755 g ml<sup>-1</sup>. For each OTU, we calculate logarithmic fold change (LFC) and corresponding standard error for enrichment in high density fractions of <sup>13</sup>C treatments relative to control. Subsequently, a one-sided Wald test was used to assess the statistical significance of LFC values with the null hypothesis that LFC was less than one standard deviation above the mean of all LFC values. We independently filtered OTUs prior to multiple comparison corrections on the basis of sparsity eliminating OTUs that failed to appear in at least 45% of high density fractions for a given comparison. P-values were adjusted for multiple comparisons using the Benjamini and Hochberg method [94]. We selected a false discovery rate of 10% to denote statistical significance.

See SI for additional information on experimental and analytical methods.

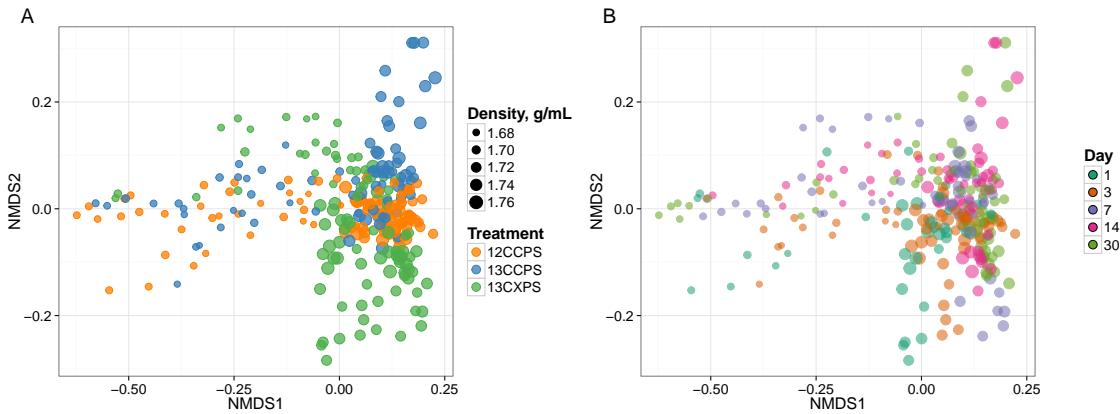
**Acknowledgements.** The authors would like to acknowledge the assistance of John Christian Gaby and Mallory Choudoir in developing the method

- used to produce  $^{13}\text{C}$ -labeled cellulose. We would also like to thank Steve Zinder, Nelson Hairston, and Nick Youngblut for providing comments that were helpful in the development of this manuscript.
- This material is based upon work supported by the Department of Energy Office of Science, Office of Biological & Environmental Research Genomic Science Program under Award Numbers DE-SC0004486 and DE-SC0010558.
- ## References
- Amundson R (2001) The carbon budget in soils. *Annu Rev Earth Planet Sci* 29(1): 535–562.
  - Batjes N-H (1996) Total carbon and nitrogen in the soils of the world. *Eur J Soil Sci* 47(2): 151–163.
  - Chapin F (2002) Principles of terrestrial ecosystem ecology. (Springer, New York).
  - Allison S-D, Wallenstein M-D, Bradford M-A (2010) Soil-carbon response to warming dependent on microbial physiology. *Nat Geosci* 3(5): 336–340.
  - Six J, Frey S-D, Thier R-K, Batten K-M (2006) Bacterial and fungal contributions to carbon sequestration in agroecosystems. *Soil Sci Soc Am J* 70(2): 555.
  - Treseder K-K, Balser T-C, Bradford M-A, Brodie E-L, Dumbinsky E-A, Eviner V-T, et al. (2011) Integrating microbial ecology into ecosystem models: challenges and priorities. *Biogeochemistry* 109(1–3): 7–18.
  - Bradford M-A, Fierer N, Reynolds J-F (2008) Soil carbon stocks in experimental mesocosms are dependent on the rate of labile carbon, nitrogen and phosphorus inputs to soils. *Funct Ecol* 22(6): 964–974.
  - Neff J-C, Asner G-P (2001) Dissolved organic carbon in terrestrial ecosystems: synthesis and a model. *Ecosystems* 4(1): 29–48.
  - McGuire K-L, Treseder K-K (2010) Microbial communities and their relevance for ecosystem models: Decomposition as a case study. *Soil Biol Biochem* 42(4): 529–535.
  - Wieder W-R, Bonan G-B, Allison S-D (2013) Global soil carbon projections are improved by modelling microbial processes. *Nat Clim Chang* 3(10): 909–912.
  - Lynd L-R, Weimer P-J, van Zyl W-H, Pretorius I-S (2002) Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol Mol Biol Rev* 66(3): 506–577.
  - Saha B-C (2003) Hemicellulose bioconversion. *J Ind Microbiol Biotechnol* 30(5): 279–291.
  - Spiridon I, Popa V-I (2008) Chapter 13 - Hemicelluloses: major sources, properties and applications. *Monomers, Polymers and Composites from Renewable Resources*, eds. Belgacem M-N, Gandini A (Elsevier, New York), pp 289–304.
  - Sun X-F, Xu F, Zhao H, Sun R-C, Fowler P, Baird M-S (2005) Physicochemical characterisation of residual hemicelluloses isolated with cyanamide-activated hydrogen peroxide from organosolv pre-treated wheat straw. *Bioreour Technol* 96(12): 1342–1349.
  - Bunnell K, Rich A, Luckett C, Wang Y-J, Martin E, Carrier D-J (2013) Plant maturity effects on the physicochemical properties and dilute acid hydrolysis of switchgrass (*Panicum virgatum*, l.) hemicelluloses. *ACS Sustain Chem Eng* 1(6): 649–654.
  - Garrett S-D (1951) Ecological groups of soil fungi: a survey of substrate relationships. *New Phytol* 50(2): 149–166.
  - Alexander M (1964) Biochemical ecology of soil microorganisms. *Annual Rev Microbiol* 18(1): 217–250.
  - Engelking B, Flessa H, Joergensen R-G (2007) Microbial use of maize cellulose and sugarcane sucrose monitored by changes in the  $^{13}\text{C}/^{12}\text{C}$  ratio. *Soil Biol Biochem* 39(8): 1888–1896.
  - Hu S, van Bruggen A-HC (1997) Microbial dynamics associated with multiphasic decomposition of  $^{14}\text{C}$ -labeled cellulose in soil. *Microb Ecol* 33(2): 134–143.
  - Rui J, Peng J, Lu Y (2009) Succession of bacterial populations during plant residue decomposition in rice field soil. *Appl Environ Microbiol* 75(14): 4879–4886.
  - Kjöller A-H, Struwe S (2002) Fungal communities, succession, enzymes, and decomposition. *Enzymes in the Environment* (Marcel Dekker, New York), pp 306–324.
  - Bastian F, Bouziri L, Nicolardot B, Ranjard L (2009) Impact of wheat straw decomposition on successional patterns of soil microbial community structure. *Soil Biol Biochem* 41(2): 262–275.
  - Garrett S-D (1963) Soil Fungi and soil fertility. (Pergamon Press, New York).
  - Bremer E, Kuikman P (1994) Microbial utilization of  $^{14}\text{C}[\text{U}]$ glucose in soil is affected by the amount and timing of glucose additions. *Soil Biol Biochem* 26(4): 511–517.
  - Frankland J-C (1998) Fungal succession – unravelling the unpredictable. *Mycol Res* 102(1): 1–15.
  - Osano T (2005) Colonization and succession of fungi during decomposition of Swida controversa leaf litter. *Mycologia* 97(3): 589–597.
  - Coleman D-C, Crossley D-A (1996) Fundamentals of soil ecology. (Academic Press, Waltham, Massachusetts).
  - Nannipieri P, Ascher J, Ceccherini M-T, Landi L, Pietramellara G, Renella G (2003) Microbial diversity and soil functions. *Eur J Soil Sci* 54(4): 655–670.
  - Strickland M-S, Lauber C, Fierer N, Bradford M-A (2009) Testing the functional significance of microbial community composition. *Ecology* 90(2): 441–451.
  - Zak D-R, Blackwood C-B, Waldrop M-P (2006) A molecular dawn for biogeochemistry. *Trends Ecol Evol* 21(6): 288–295.
  - Reed H-E, Martiny J-BH (2007) Testing the functional significance of microbial composition in natural communities. *FEMS Microbiol Ecol* 62(2): 161–170.
  - Schimel J-P, Schaeffer S-M (2012) Microbial control over carbon cycling in soil. *Front Microbiol* 3: 348. doi: 10.3389/fmicb.2012.00348
  - Allison S-D, Martiny J-BH (2008) Resistance resilience, and redundancy in microbial communities. *Proc Natl Acad Sci USA* 105(Supplement 1): 11512–11519.
  - Cavigelli M-A, Robertson G-P (2000) The functional significance of denitrifier community composition in a terrestrial ecosystem. *Ecology* 81(5): 1402–1414.
  - Carney K-M, Matson P-A, Bohannan B-JM (2004) Diversity and composition of tropical soil nitrifiers across a plant diversity gradient and among land-use types. *Ecol Lett* 7(8): 684–694.
  - Hawkes C-V, Wren I-F, Herman D-J, Firestone M-K (2005) Plant invasion alters nitrogen cycling by modifying the soil nitrifying community. *Ecol Lett* 8(9): 976–985.
  - Webster G, Embley T-M, Freitag T-E, Smith Z, Prosser J-I (2005) Links between ammonia oxidizer species composition, functional diversity and nitrification kinetics in grassland soils. *Environ Microbiol* 7(5): 676–684.
  - Gulledge J, Doyle A-P, Schimel J-P (1997) Different  $\text{NH}_4^+$ -inhibition patterns of soil  $\text{CH}_4$  consumption: A result of distinct  $\text{CH}_4$ -oxidizer populations across sites? *Soil Biol Biochem* 29(1): 13–21.
  - Hsu S-F, Buckley D-H (2009) Evidence for the functional significance of diazotroph community structure in soil. *ISME J* 3(1): 124–136.
  - Janssen P-H (2006) Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl Environ Microbiol* 72(3): 1719–1728.
  - Buckley D-H, Schmidt T-M (2002) Exploring the diversity of soil – a microbial rainforest. *Biodiversity of Microbial Life: Foundation of Earth’s Biosphere*, ed. Reysenbach A-L (Wiley, New York, New York), pp 183–208.
  - Chen Y, Murrell J-C (2010) When metagenomics meets stable-isotope probing: progress and perspectives. *Trends Microbiol* 18(4): 157–163.
  - Radajewski S, Ineson P, Parekh N-R, Murrell J-C (2000) Stable-isotope probing as a tool in microbial ecology. *Nature* 403(6770): 646–649.
  - Manefield M, Whitley A-S, Griffiths R-I, Bailey M-J (2002) RNA stable isotope probing: a novel means of linking microbial community function to phylogeny. *Appl Environ Microbiol* 68(11): 5367–5373.
  - McDonald I-R, Radajewski S, Murrell J-C (2005) Stable isotope probing of nucleic acids in methanotrophs and methylotrophs: A review. *Org Geochem* 36(5): 779–787.
  - Morris S-A, Radajewski S, Willison T-W, Murrell J-C (2002) Identification of the functionally active methanotroph population in a peat soil microcosm by stable-isotope probing. *Appl Environ Microbiol* 68(3): 1446–1453.
  - Hutchens E, Radajewski S, Dumont M-G, McDonald I-R, Murrell J-C (2004) Analysis of methanotrophic bacteria in Movie Cave by stable isotope probing. *Environ Microbiol* 6(2): 111–120.

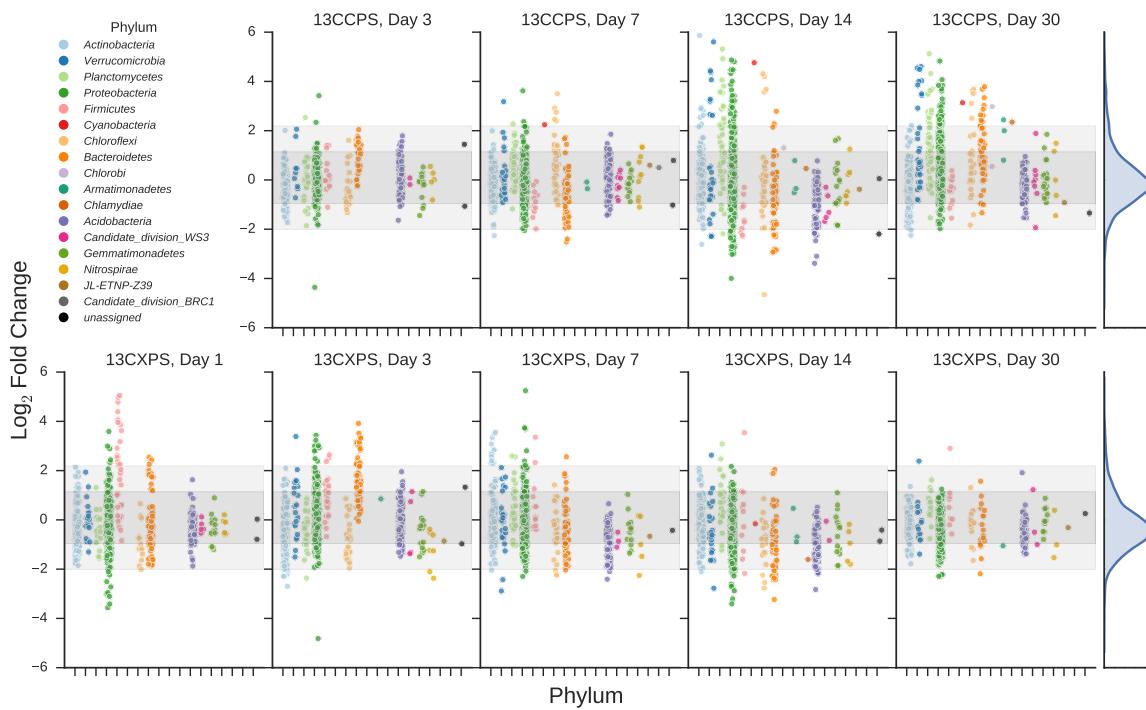
48. Lueders T, Manefield M, Friedrich M (2004) Enhanced sensitivity of DNA- and rRNA-based stable isotope probing<sup>1125</sup> by fractionation and quantitative analysis of isopycnic centrifugation gradients.. *Environ Microbiol* 6: 73–8.
- 1045 49. DeRito C-M, Pumphrey G-M, Madsen E-L (2005) Use of field-based stable isotope probing to identify adapted populations and track carbon flow through a phenol-degrading<sup>1130</sup> soil microbial community. *Appl Environ Microbiol* 71(12): 7858–7865.
- 1050 50. Ziegler S-E, White P-M, Wolf D-C, Thoma G-J (2005) Tracking the fate and recycling of <sup>13</sup>C-labeled glucose in soil. *Soil Sci* 170(10): 767–778.<sup>1135</sup>
- 1055 51. Buckley D-H, Huangyutitham V, Hsu S-F, Nelson T-A (2007) Stable isotope probing with <sup>15</sup>N achieved by disentangling the effects of genome G+C content and isotope enrichment on DNA density. *Appl Environ Microbiol* 73(10): 3189–3195.<sup>1140</sup>
- 1060 52. Birnie G-D (1978) Centrifugal separations in Molecular and cell biology. (Butterworth & Co Publishers Ltd, Boston).
- 1065 53. Holben W-E, Harris D (1995) DNA-based monitoring of total bacterial community structure in environmental samples. *Molecular Ecology* 4(5): 627–632.<sup>1145</sup>
- 1070 54. Nüsslein K, Tiehd J-M (1999) Soil bacterial community shift correlated with change from forest to pasture vegetation in a tropical soil. *Appl Environ Microbiol* 65(8): 3622–3626.
- 1075 55. Lu Y, Conrad R (2005) In situ stable isotope probing<sup>1150</sup> of methanogenic archaea in the rice rhizosphere. *Science* 309(5737): 1088–1090.
- 1080 56. Lueders T, Pommernke B, Friedrich M-W (2004) Stable-isotope probing of microorganisms thriving at thermodynamic limits: syntrophic propionate oxidation in flooded<sup>1155</sup> soil. *Appl Environ Microbiol* 70(10): 5778–5786.
- 1085 57. DeRito C-M, Pumphrey G-M, Madsen E-L (2005) Use of field-based stable isotope probing to identify adapted populations and track carbon flow through a phenol-degrading<sup>1160</sup> soil microbial community. *Appl Environ Microbiol* 71(12): 7858–7865.
- 1090 58. Aoyagi T, Hanada S, Itoh H, Sato Y, Ogata A, Friedrich M-W, et al. (2015) Ultra-high-sensitivity stable-isotope probing of rRNA by high-throughput sequencing of isopycnic centrifugation gradients. *Environ Microbiol Rep* 7(2): 282–287.<sup>1165</sup>
- 1095 59. Verastegui Y, Cheng J, Engel K, Kolczynski D, Mortimer S, Lavigne J, et al. (2014) Multisubstrate isotope labeling and metagenomic analysis of active soil bacterial communities. *mBio* 5(4): e01157–14.<sup>1170</sup>
- 1100 60. Anderson M-J (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 26(1): 32–46.
- 1105 61. Kembel S-W, Wu M, Eisen J-A, Green J-L (2012) Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. *PLoS Comput Biol* 8(10): e1002743.<sup>1175</sup>
- 1110 62. Klappenbach J-A, Dunbar J-M, Schmidt T-M (2000) rRNA Operon copy number reflects ecological strategies of bacteria. *Appl Environ Microbiol* 66(4): 1328–1333.<sup>1180</sup>
- 1115 63. Webb C-O (2000) Exploring the phylogenetic structure of ecological communities: an example for rain forest trees.. *Am Nat* 156(2): 145–155.
- 1120 64. Martiny A-C, Treseder K, Pusch G (2013) Phylogenetic conservatism of functional traits in microorganisms. *ISME J* 7(4): 830–838.
- 1125 65. Evans S-E, Wallenstein M-D (2014) Climate change alters ecological strategies of soil bacteria. *Ecol Lett* 17(2): 155–164.
- 1130 66. Berlemont R, Martiny A-C (2013) Phylogenetic distribution of potential cellulases in bacteria. *Appl Environ Microbiol* 79(5): 1545–1554.<sup>1190</sup>
- 1135 67. Fierer N, Bradford M-A, Jackson R-B (2007) Toward an ecological classification of soil bacteria. *Ecology* 88(6): 1354–1364.
- 1140 68. Bergmann G-T, Bates S-T, Eilers K-G, Lauber C-L, Caporaso J-G, Walters W-A, et al. (2011) The under-recognized dominance of *Verrucomicrobia* in soil bacterial communities. *Soil Biol Biochem* 43(7): 1450–1455.
- 1145 69. Buckley D-H, Schmidt T-M (2001) Environmental factors influencing the distribution of rRNA from *Verrucomicrobia* in soil. *FEMS Microbiol Ecol* 35(1): 105–112.<sup>1200</sup>
- 1150 70. Wertz J-T, Kim E, Breznak J-A, Schmidt T-M, Rodrigues J-LM (2011) Genomic and physiological characterization of the *Verrucomicrobia* isolate *Diplosphaera colitermitum* gen. nov. sp. nov., reveals microaerophily and nitrogen fixation genes. *Appl Environ Microbiol* 78(5): 1544–1555.
71. Otsuka S, Ueda H, Suenaga T, Uchino Y, Hamada M, Yokota A, et al. (2012) *Roseimicrobium gellanolyticum* gen. nov. sp. nov., a new member of the class *Verrucomicrobiae*. *Int J Syst Evol Microbiol* 63(Pt 6): 1982–1986.
72. Fierer N, Ladau J, Clemente J-C, Leff J-W, Owens S-M, Pollard K-S, et al. (2013) Reconstructing the microbial diversity and function of pre-agricultural tallgrass prairie soils in the united states. *Science* 342(6158): 621–624.
73. Chin K-J, Hahn D, Hengstmann U, Liesack W, Janssen P-H (1999) Characterization and identification of numerically abundant culturable bacteria from the anoxic bulk soil of rice paddy microcosms. *Appl Environ Microbiol* 65(11): 5042–5049.
74. Herlemann D-PR, Lundin D, Labrenz M, Jurgens K, Zheng Z, Aspeborg H, et al. (2013) Metagenomic de novo assembly of an aquatic representative of the verrucomicrobial class *Spartobacteria*. *mBio* 4(3): e0056912.
75. Schellenberger S, Kolb S, Drake H-L (2010) Metabolic responses of novel cellulolytic and saccharolytic agricultural soil Bacteria to oxygen. *Environ Microbiol* 12(4): 845–861.
76. Zarda B, Hahn D, Chatzinotas A, Schnhuber W, Neef A, Amann R-I, et al. (1997) Analysis of bacterial community structure in bulk soil by in situ hybridization. *Arch Microbiol* 168(3): 185–192.
77. Chatzinotas A, Sandaa R-A, Schnhuber W, Amann R, Daee F-L, Torsvik V, et al. (1998) Analysis of broad-scale differences in microbial community composition of two pristine forest soils. *Syst Appl Microbiol* 21(4): 579–587.
78. Buckley D-H, Schmidt T-M (2003) Diversity and dynamics of microbial communities in soils from agro-ecosystems. *Environ Microbiol* 5(6): 441–452.
79. Hug L-A, Castelle C-J, Wrighton K-C, Thomas B-C, Sharon I, Frischkorn K-R, et al. (2013) Community genomic analyses constrain the distribution of metabolic traits across the *Chloroflexi* phylum and indicate roles in sediment carbon cycling. *Microbiome* 1(1): 22.
80. Goldfarb K-C, Karaoz U, Hanson C-A, Santee C-A, Bradford M-A, Treseder K-K, et al. (2011) Differential growth responses of soil bacterial taxa to carbon substrates of varying chemical recalcitrance. *Front Microbiol* 2: 94. doi: 10.3389/fmicb.2011.00094
81. Cole J-K, Gieler B-A, Heisler D-L, Palisoc M-M, Williams A-J, Dohnalkova A-C, et al. (2013) *Kallotenuus paprylyticum* gen. nov. sp. nov., a cellulolytic and filamentous thermophile that represents a novel lineage (*Kallotenuales* ord. nov., *Kallotenuaceae* fam. nov.) within the class *Chloroflexia*. *Int J Syst Evol Microbiol* 63(Pt 12): 4675–4682.
82. Rienzi S-CD, Sharon I, Wrighton K-C, Koren O, Hug L-A, Thomas B-C, et al. (2013) The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to *Cyanobacteria*. *eLIFE* 2: e01102.
83. Gromov B-V, Mamkaeva K-A (1972) Electron microscopic study of parasitism by *Bdellovibrio chlorellavorus* bacteria on cells of the green alga *Chlorella vulgaris*. *Tsitologiya* 14(2): 256–260.
84. Cleveland C-C, Nemergut D-R, Schmidt S-K, Townsend A-R (2007) Increases in soil respiration following labile carbon additions linked to rapid shifts in soil microbial community composition. *Biogeochemistry* 82(3): 229–240.
85. Casida L-E (1983) Interaction of *Agromyces ramosus* with other bacteria in soil. *Appl Environ Microbiol* 46(4): 881–888.
86. Lueders T, Kindler R, Miltner A, Friedrich M-W, Kaestner M (2006) Identification of bacterial micropredators distinctively active in a soil microbial food web. *Appl Environ Microbiol* 72(8): 5342–5348.
87. Moore J-C, Walter D-E, Hunt H-W (1988) Arthropod regulation of micro- and mesobiota in below-ground detrital food webs. *Annu Rev Entomol* 33(1): 419–435.
88. Kaiser C, Franklin O, Dieckmann U, Richter A (2014) Microbial community dynamics alleviate stoichiometric constraints during litter decay. *Ecol Lett* 17(6): 680–690.
89. Wieder W-R, Grandy A-S, Kallenbach C-M, Bonan G-B (2014) Integrating microbial physiology and physicochemical principles in soils with the Microbial-MIneral Carbon Stabilization (MIMICS) model. *Biogeosciences* 11(14): 3899–3917.
90. Schimel J (1995) Ecosystem consequences of microbial diversity and community structure. *Arctic and alpine biodiversity: patterns, causes and ecosystem consequences*. eds. Chapin III F-S, Körner C, Ecological Studies



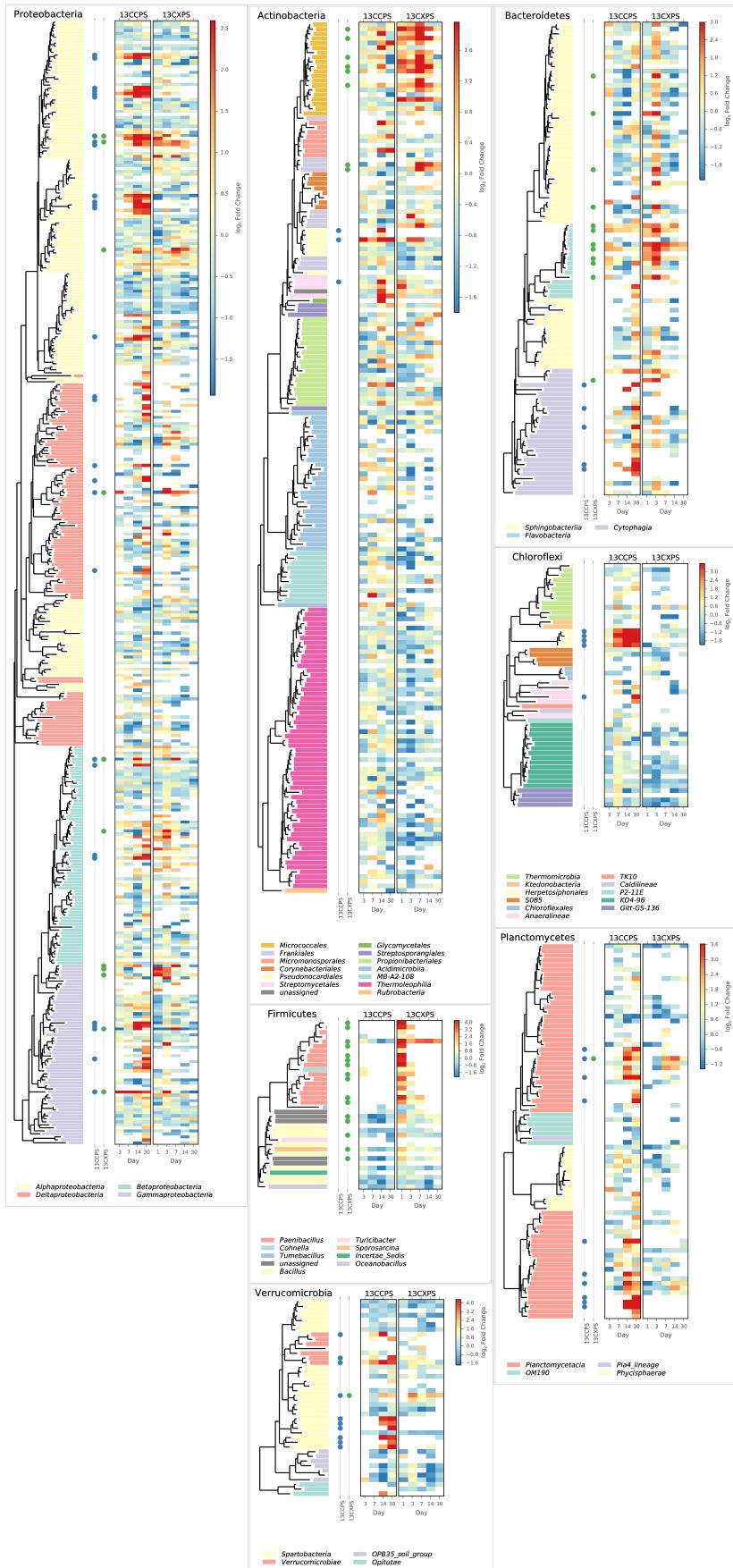
- (Springer, Berlin Heidelberg), pp 239–254.
91. Schneckenberger K, Demin D, Stahr K, Kuzyakov Y (2008) Microbial utilization and mineralization of  $^{14}\text{C}$  glucose added in six orders of concentration to soil. *Soil Biol Biochem* 40(8): 1981–1988.
- 1210 92. Love M-I, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12): 550.
- 1215 93. McMurdie P-J, Holmes S (2014) Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol* 10(4): e1003531.
94. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57(1): 289–300.



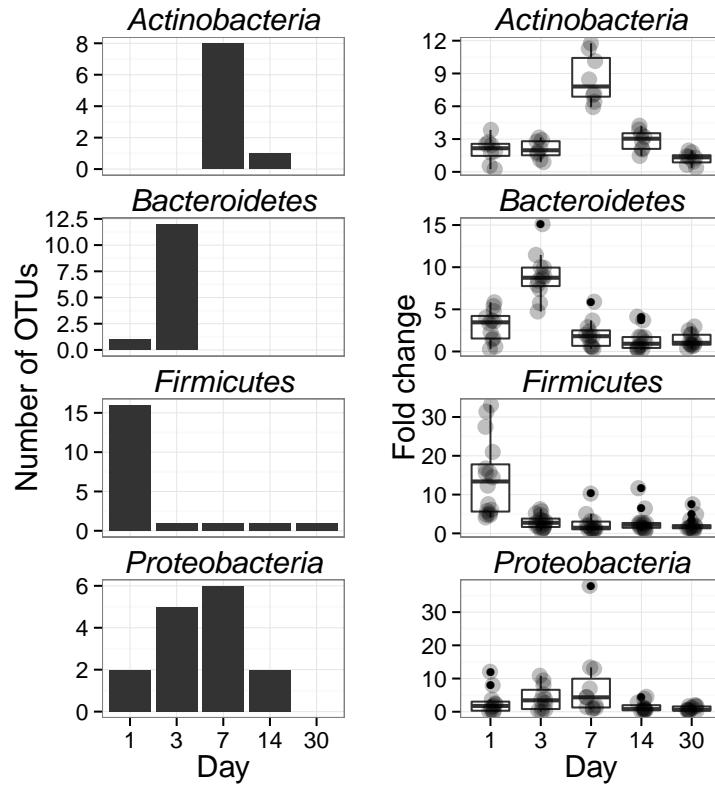
**Fig. 1.** NMDS ordination of SSU rRNA gene sequence composition in gradient fractions shows that it is a function of many factors including fraction density, isotopic labeling, and time. Dissimilarity in SSU rRNA gene sequence composition was quantified using the weighted UniFrac metric. SSU rRNA gene sequences were surveyed in twenty gradient fractions at each sampling point for each treatment (Figure S1).  $^{13}\text{C}$ -labeling of DNA is apparent because the SSU rRNA gene sequence composition of gradient fractions from  $^{13}\text{C}$  and control treatments differ at high density. Each point on the NMDS plot represents one gradient fraction. SSU rRNA gene sequence composition differences between gradient fractions were quantified by the weighted Unifrac metric. The size of each point is positively correlated with density and colors indicate the treatment (A) or day (B).



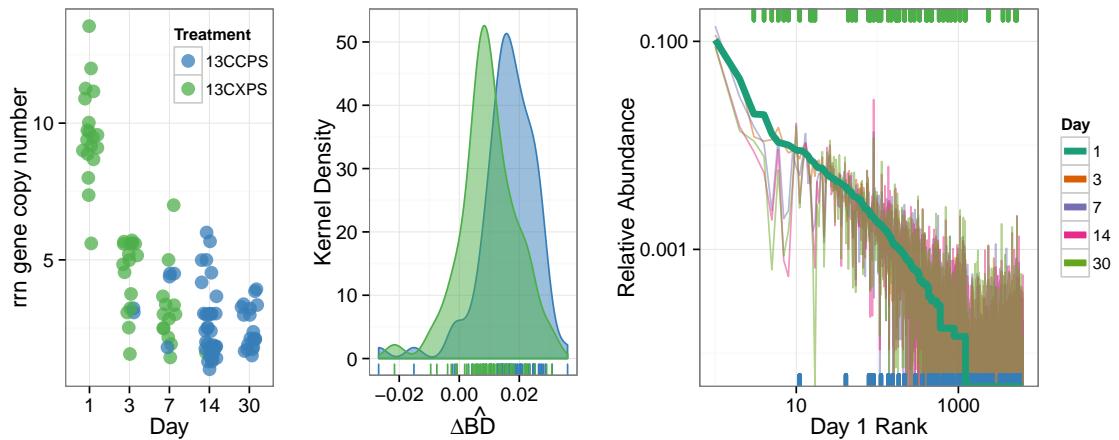
**Fig. 2.** Enrichment of OTUs in either  $^{13}\text{C}$ -cellulose (13CCPS, upper panels) or  $^{13}\text{C}$ -xylose (13CXPS, bottom panels) treatments relative to control, expressed as LFC (see Methods). Each point indicates the LFC for a single OTU. High enrichment values indicate an OTU is likely  $^{13}\text{C}$ -labeled. Different colors represent different phyla and different panels represent different days. The final column shows the frequency distribution of LFC values in each row. Within each panel, shaded areas are used to indicate one standard deviation (dark shading) or two standard deviations (light shading) about the mean of all LFC values.



**Fig. 3.** Phylogenetic position of cellulose responders and xylose responders in the context of all OTUs that passed sparsity independent filtering criteria (see Methods). Only those phyla that contain responders are shown. Colored dots are used to identify xylose responders (green) and cellulose responders (blue). The heatmaps indicate enrichment in high density fractions relative to control (represented as LFC) for each OTU in response to both  $^{13}\text{C}$ -cellulose (13CCPS, leftmost heatmap) and  $^{13}\text{C}$ -xylose (13CXPS, rightmost heatmap) with values for different days in each heatmap column. High enrichment values (represented as LFC) provide evidence of  $^{13}\text{C}$ -labeled DNA.

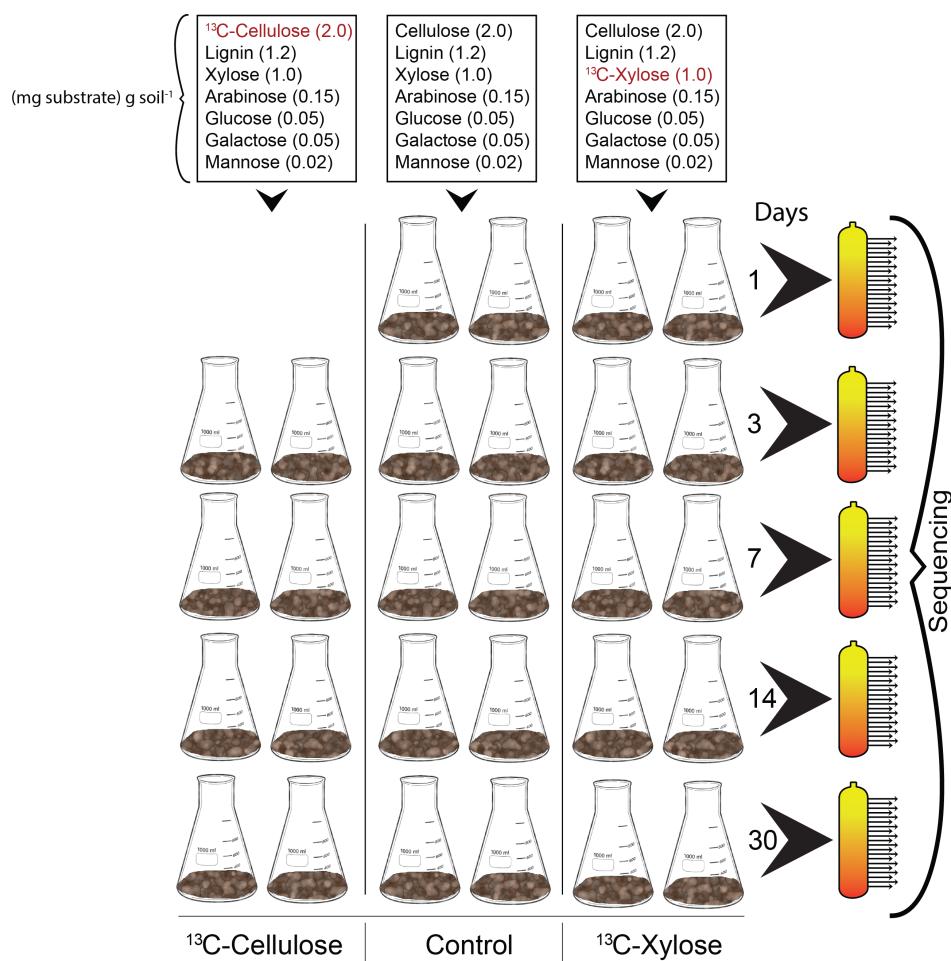


**Fig. 4.** Xylose responders in the *Actinobacteria*, *Bacteroidetes*, *Firmicutes* exhibit distinct temporal dynamics of  $^{13}\text{C}$ -labeling. The left column shows counts of  $^{13}\text{C}$ -xylose responders in the *Actinobacteria*, *Bacteroidetes*, *Firmicutes* and *Proteobacteria* at days 1, 3, 7 and 30. The right panel shows OTU enrichment in high density gradient fractions (gray points, expressed as fold change) for responders as well as a boxplot for the distribution of fold change values (The box extends one interquartile range, whiskers extend 1.5 times the IR, and small dots are outliers (i.e. beyond 1.5 times the IR)). Each day in the right column shows all responders (i.e. OTUs that responded to xylose at any point in time). High enrichment values indicates OTU DNA is likely  $^{13}\text{C}$ -labeled.

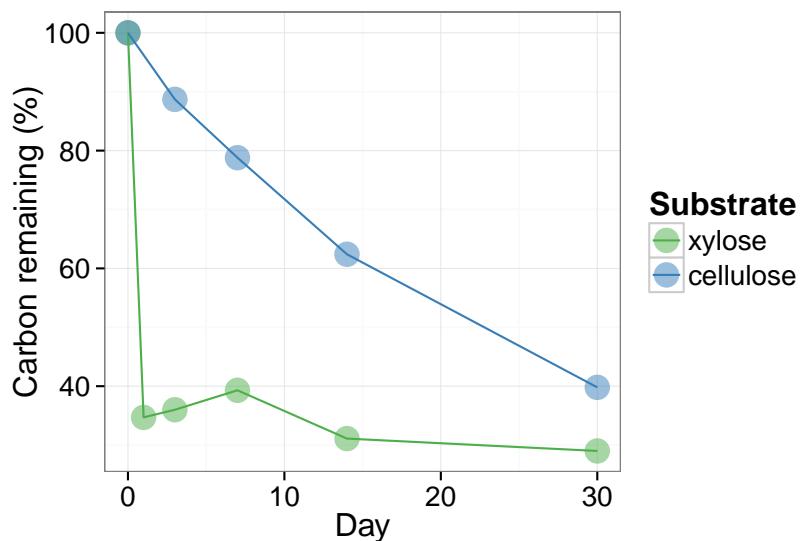


**Fig. 5.** Characteristics of xylose responders (green) and cellulose responders (blue) based on estimated *rrn* copy number (A),  $\Delta\text{BD}$  (B), and relative abundance in non-fractionated DNA (C). The estimated *rrn* copy number of all responders is shown versus time (A). Kernel density histogram of  $\Delta\text{BD}$  values shows cellulose responders had higher average  $\Delta\text{BD}$  than xylose responders indicating higher average atom %  $^{13}\text{C}$  in OTU DNA (B). The final panel indicates the rank relative abundance of all OTUs observed in the non-fractionated DNA (C) where rank was determined at day 1 (bold line) and relative abundance for each OTU is indicated for all days by colored lines (see legend). Xylose responders (green ticks) have higher relative abundance in non-fractionated DNA than cellulose responders (ticks are based on day 1 relative abundance).

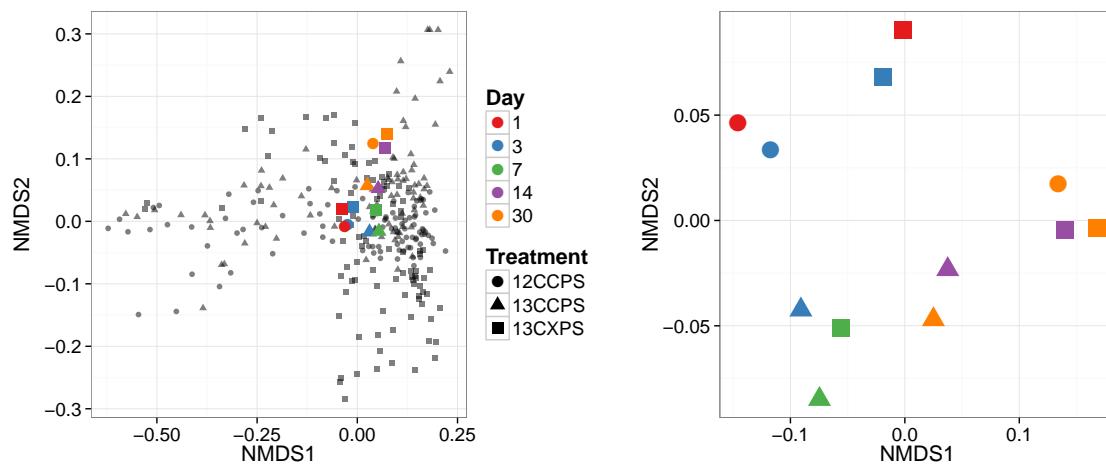
**Supplemental Figures and Tables**



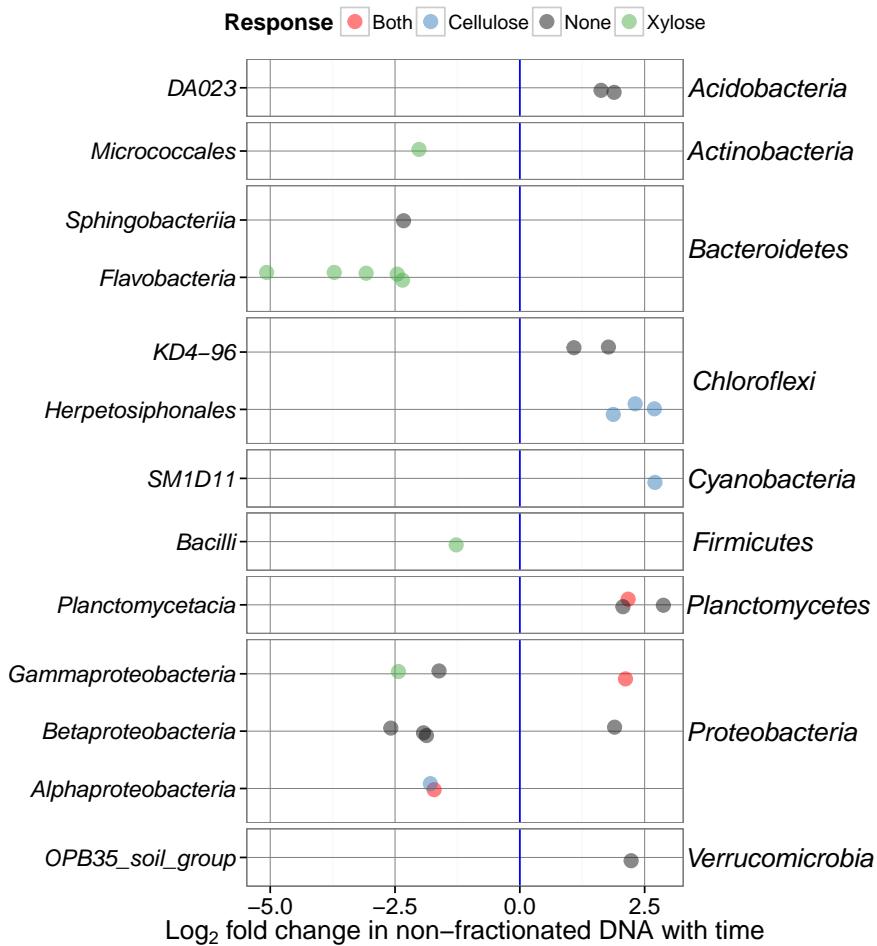
**Fig. S1.** We added a carbon mixture with inorganic salts and amino acids (not shown here) to each soil microcosm where the only difference between treatments was the <sup>13</sup>C-labeled isotope (in red). At days 1, 3, 7, 14, and 30 replicate microcosms were destructively harvested for downstream molecular applications. DNA from each treatment and time ( $n = 14$ ) was subjected to CsCl density gradient centrifugation and density gradients were fractionated (orange tubes wherein each arrow represents a fraction from the density gradient). SSU rRNA genes from each gradient fraction were PCR amplified and sequenced. In addition, SSU rRNA genes were also PCR amplified and sequenced from non-fractionated DNA to represent the soil microbial community.



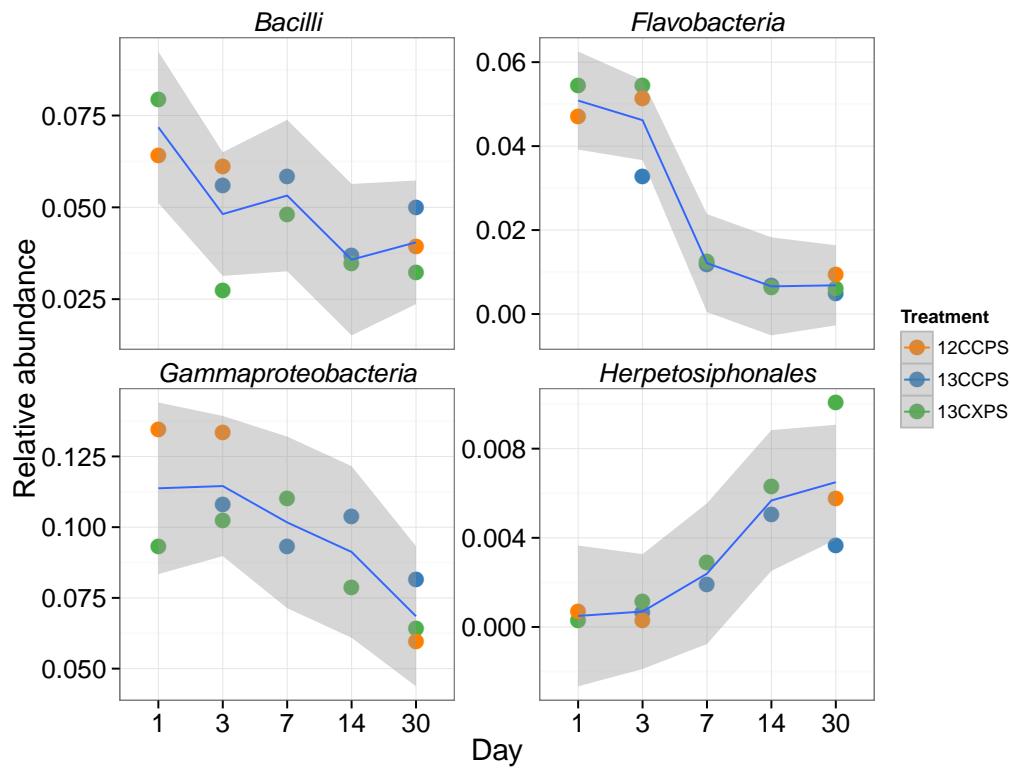
**Fig. S2.** The metabolism of <sup>13</sup>C-xylose and <sup>13</sup>C-cellulose is indicated by the percentage of the added <sup>13</sup>C that remains in soil over time.



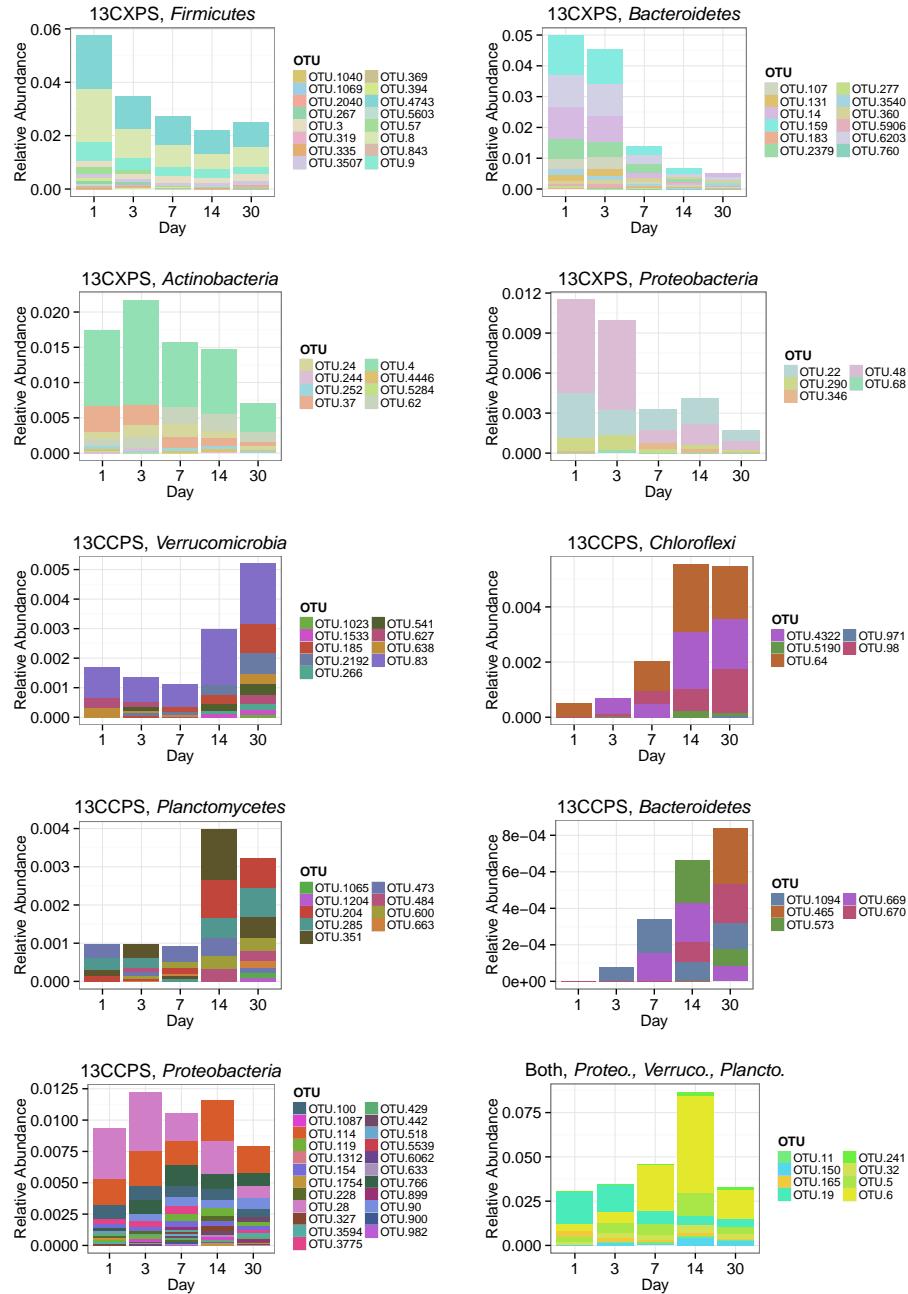
**Fig. S3.** NMDS analysis of SSU rRNA gene composition in non-fractionated DNA indicates that isotopic labelling does not alter overall microbial community composition, microbial community composition in the soil microcosms changes over time, and variance in non-fractionated DNA is smaller than variance in fractionated DNA. SSU rRNA gene sequences were determined for non-fractionated DNA from the unlabeled control,  $^{13}\text{C}$ -xylose, and  $^{13}\text{C}$ -cellulose treatments over time (colors indicate time, different symbols used for different treatments). Distance in SSU rRNA gene composition was quantified with the UniFrac metric. The leftmost panel indicates NMDS of data from both non-fractionated and fractionated samples. The rightmost panel indicates NMDS of data only from non-fractionated DNA. Statistical analysis is presented in main text.



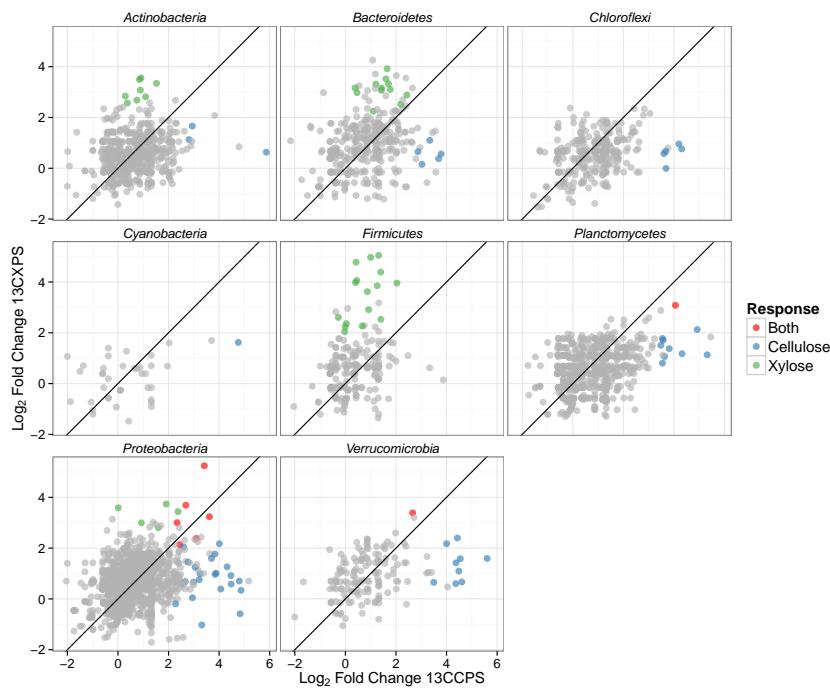
**Fig. S4.** Change in non-fractionated DNA relative abundance versus time (expressed as LFC) for OTUs that changed significantly ( $P\text{-value} < 0.10$ , Wald test). Each panel shows one phylum (labeled on the right). The taxonomic class is indicated on the left. Colors represent results shown in Figure 2 and Figure 3. OTUs that responded to just xylose are shown in green, just cellulose in blue, and both xylose and cellulose in red.



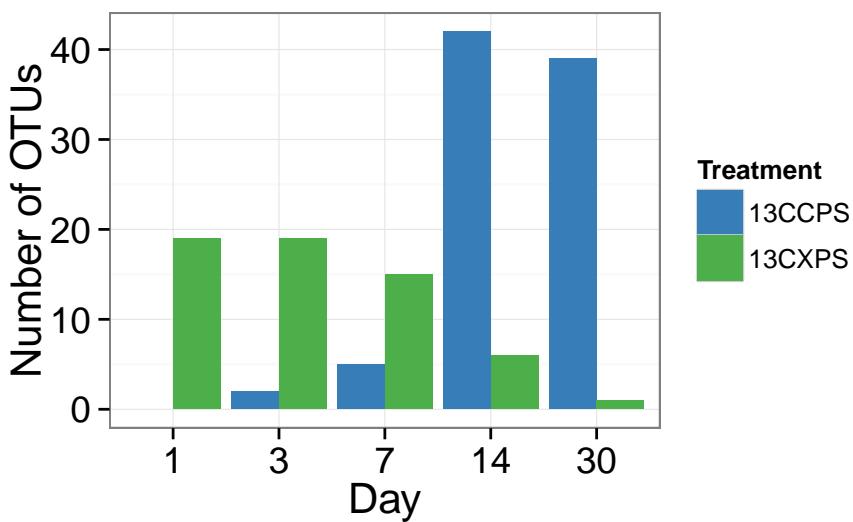
**Fig. S5.** Relative abundance in non-fractionated DNA versus time for classes that changed significantly. Samples from different treatments are labeled with different colors as indicated in the scale. Statistical analysis is presented in main text.



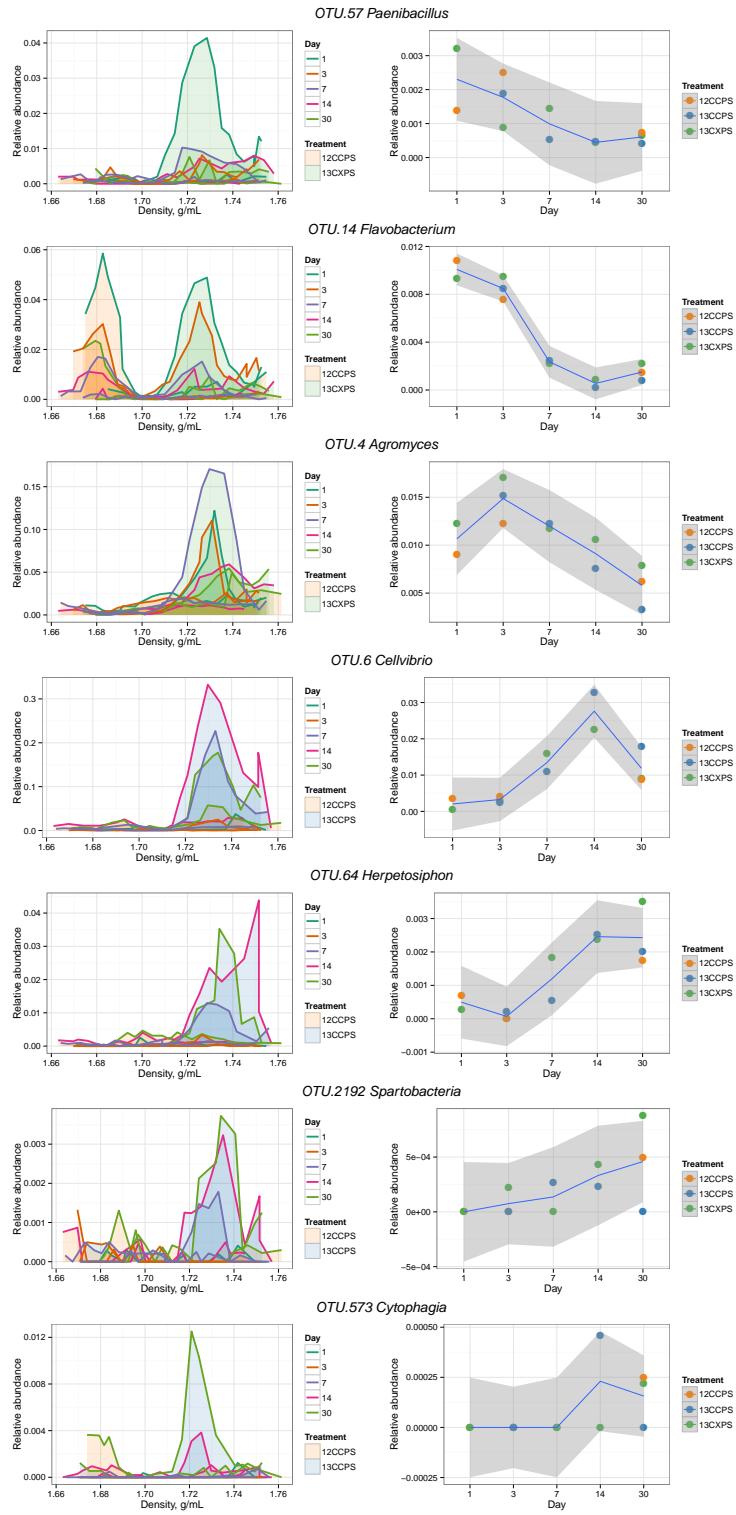
**Fig. S6.** Change in relative abundance in non-fractionated DNA over time for xylose responders (13CXPS) and cellulose responders (13CCPS). Each panel represents a responders to the indicated substrate (i.e. cellulose (13CCPS) or xylose (13CXPS)) within the indicated phylum except for the lower right panel which shows all responders to both xylose and cellulose. The abbreviations Proteo., Verruco., and Plancto., correspond to *Proteobacteria*, *Verrucomicrobia*, and *Planctomycetes*, respectively.



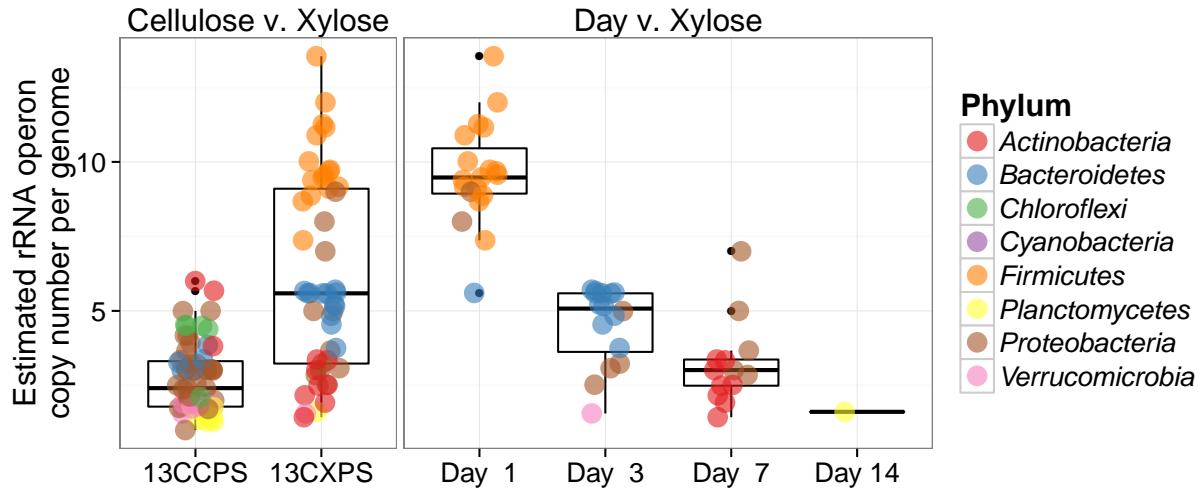
**Fig. S7.** Maximum enrichment at any point in time in high density fractions of  $^{13}\text{C}$ -treatments relative to control (expressed as LFC) shown for  $^{13}\text{C}$ -cellulose versus  $^{13}\text{C}$ -xylose treatments. Each point represents an OTU. Blue points are cellulose responders, green xylose responders, red are responders to both xylose and cellulose, and gray points are OTUs that did not respond to either substrate. Line indicates a slope of one.



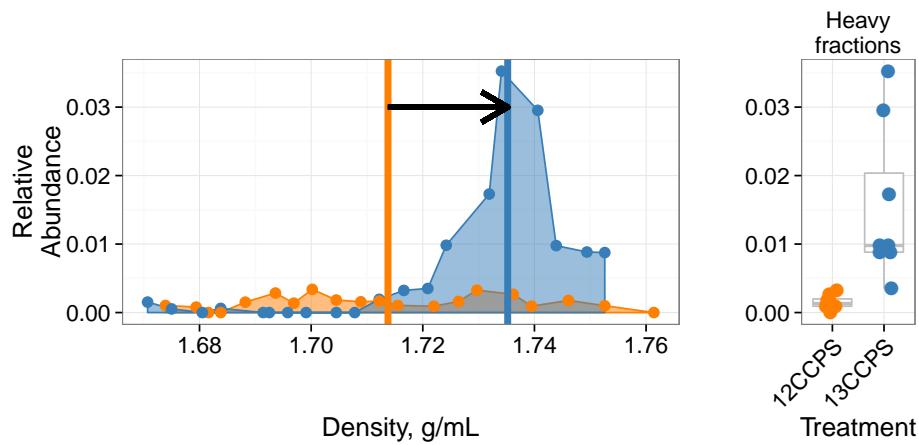
**Fig. S8.** Counts of xylose responders and cellulose responders over time.



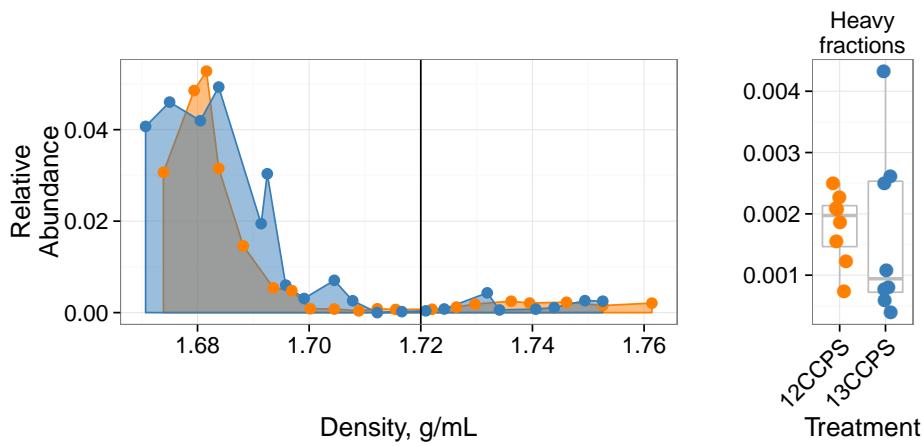
**Fig. S9.** Raw data from individual responders highlighted in the main text (see Results). The left column shows OTU relative abundance in gradient density fractions for the indicated treatment pair at each sampling point. Time is indicated by the line color (see legend). Gradient profiles are shaded to represent the different treatments where orange represents “control”, blue “<sup>13</sup>C-cellulose”, and green “<sup>13</sup>C-xylose.” The right column shows the relative abundance of each OTU in non-fractionated DNA. Enrichment in the high density fractions of <sup>13</sup>C-treatments indicates an OTU likely has <sup>13</sup>C-labeled DNA.



**Fig. S10.** Estimated *rRNA* copy number for xylose and cellulose responders. The leftmost panel contrasts estimated *rRNA* copy number for cellulose (13CCPS) and xylose (13CXPS) responders. The right panel shows estimated *rRNA* copy number versus time of first response for xylose responders. Colors denote the phylum of the OTUs (see legend).



**Fig. S11.** Density profile for a single cellulose responder in the  $^{13}\text{C}$ -cellulose treatment (blue) and control (orange). Vertical lines show center of mass for each density profile and the arrow denotes the magnitude and direction of  $\Delta BD$ . Right panel shows relative abundance values in the high density fractions (The boxplot line is the median value. The box spans one interquartile range (IR) about the median, whiskers extend 1.5 times the IR, and the dots indicate outlier values beyond 1.5 times the IR).



**Fig. S12.** Density profile for a single non-responder OTU. The  $^{13}\text{C}$ -cellulose treatment is in blue and the control treatment is in orange. The vertical line shows where high density fractions begin as defined in our analysis. The right panel shows relative abundance values in the high density fractions for each gradient (The boxplot line is the median value. The box spans one interquartile range (IR) about the median, whiskers extend 1.5 times the IR and the dots indicate outlier values beyond 1.5 times the IR).

Table S1:  $^{13}\text{C}$ -xylose responders BLAST against Living Tree Project

OTU ID	Fold change <sup>a</sup>	Day <sup>b</sup>	All days <sup>c</sup>	Top BLAST hits	BLAST %ID	Phylum;Class;Order
OTU.1040	4.78	1	1	<i>Paenibacillus daejeonensis</i>	100.0	<i>Firmicutes Bacilli Bacillales</i>
OTU.1069	3.85	1	1	<i>Paenibacillus terrigena</i>	100.0	<i>Firmicutes Bacilli Bacillales</i>
OTU.107	2.25	3	3	<i>Flavobacterium sp. 15C3</i> , <i>Flavobacterium banpakuense</i>	99.54	<i>Bacteroidetes Flavobacteria Flavobacteriales</i>
OTU.11	5.25	7	7	<i>Stenotrophomonas pavani</i> i, <i>Stenotrophomonas maltophilia</i> , <i>Pseudomonas geniculata</i>	99.54	<i>Proteobacteria Gammaproteobacteria Xanthomonadales</i>
OTU.131	3.07	3	3	<i>Flavobacterium fluvii</i> , <i>Flavobacterium bacterium HMD1033</i> , <i>Flavobacterium sp. HMD1001</i>	100.0	<i>Bacteroidetes Flavobacteria Flavobacteriales</i>
OTU.14	3.92	3	1, 3	<i>Flavobacterium oncorhynchi</i> , <i>Flavobacterium glycines</i> , <i>Flavobacterium succinicans</i>	99.09	<i>Bacteroidetes Flavobacteria Flavobacteriales</i>
OTU.150	3.08	14	14	No hits of at least 90% identity	86.76	<i>Planctomycetes Planctomycetacia Planctomycetales</i>
OTU.159	3.16	3	3	<i>Flavobacterium hibernum</i>	98.17	<i>Bacteroidetes Flavobacteria Flavobacteriales</i>
OTU.165	2.38	3	3	<i>Rhizobium skieniewicense</i> , <i>Rhizobium vignae</i> , <i>Rhizobium larrymoorei</i> , <i>Rhizobium alkalisoli</i> , <i>Rhizobium galegae</i> , <i>Rhizobium huautlense</i>	100.0	<i>Proteobacteria Alphaproteobacteria Rhizobiales</i>
OTU.183	3.31	3	3	No hits of at least 90% identity	89.5	<i>Bacteroidetes Sphingobacteriia Sphingobacteriales</i>
OTU.19	2.14	7	7	<i>Rhizobium alamii</i> , <i>Rhizobium mesosinicum</i> , <i>Rhizobium mongolense</i> , <i>Arthrobacter viscosus</i> , <i>Rhizobium sullae</i> , <i>Rhizobium yanglingense</i> , <i>Rhizobium loessense</i>	99.54	<i>Proteobacteria Alphaproteobacteria Rhizobiales</i>
OTU.2040	2.91	1	1	<i>Paenibacillus pectinilyticus</i>	100.0	<i>Firmicutes Bacilli Bacillales</i>
OTU.22	2.8	7	7, 14	<i>Paracoccus sp. NB88</i>	99.09	<i>Proteobacteria Alphaproteobacteria Rhodobacterales</i>
OTU.2379	3.1	3	3	<i>Flavobacterium pectinovorum</i> , <i>Flavobacterium sp. CS100</i>	97.72	<i>Bacteroidetes Flavobacteria Flavobacteriales</i>
OTU.24	2.81	7	7	<i>Cellulomonas aerilata</i> , <i>Cellulomonas humilata</i> , <i>Cellulomonas terrae</i> , <i>Cellulomonas soli</i> , <i>Cellulomonas xylanilytica</i>	100.0	<i>Actinobacteria Micrococcales Cellulomonadaceae</i>
OTU.241	3.38	3	3, 14	No hits of at least 90% identity	87.73	<i>Verrucomicrobia Spartobacteria Chthoniobacteriales</i>
OTU.244	3.08	7	7	<i>Cellulosimicrobium funkei</i> , <i>Cellulosimicrobium terreum</i>	100.0	<i>Actinobacteria Micrococcales Promicromonosporaceae</i>
OTU.252	3.34	7	7	<i>Promicromonospora thailandica</i>	100.0	<i>Actinobacteria Micrococcales Promicromonosporaceae</i>
OTU.267	4.97	1	1	<i>Paenibacillus pabuli</i> , <i>Paenibacillus tundrae</i> , <i>Paenibacillus taichungensis</i> , <i>Paenibacillus xylohexedens</i> , <i>Paenibacillus xylanilyticus</i>	100.0	<i>Firmicutes Bacilli Bacillales</i>
OTU.277	3.52	3	3	<i>Solibius ginsengiterrae</i>	95.43	<i>Bacteroidetes Sphingobacteriia Sphingobacteriales</i>

Table S1 – continued from previous page

OTU ID	Fold change	Day	All days	Top BLAST hits	BLAST %ID	Phylum;Class;Order
OTU.290	3.59	1	1	<i>Pantoea spp.</i> , <i>Kluyvera spp.</i> , <i>Klebsiella spp.</i> , <i>Erwinia spp.</i> , <i>Enterobacter spp.</i> , <i>Buttiauxella spp.</i>	100.0	<i>Proteobacteria</i> <i>Gammaproteobacteria</i> <i>Enterobacteriales</i>
OTU.3	2.61	1	1	[ <i>Brevibacterium</i> ] <i>frigoritolerans</i> , <i>Bacillus sp. LMG 20238</i> , <i>Bacillus coahuilensis m4-4</i> , <i>Bacillus simplex</i>	100.0	<i>Firmicutes</i> <i>Bacilli</i> <i>Bacillales</i>
OTU.319	3.98	1	1	<i>Paenibacillus xinjiangensis</i>	97.25	<i>Firmicutes</i> <i>Bacilli</i> <i>Bacillales</i>
OTU.32	3.0	3	3, 7, 14	<i>Sandaracinus amyloyticus</i>	94.98	<i>Proteobacteria</i> <i>Deltaproteobacteria</i> <i>Myxococcales</i>
OTU.335	2.53	1	1	<i>Paenibacillus thailandensis</i>	98.17	<i>Firmicutes</i> <i>Bacilli</i> <i>Bacillales</i>
OTU.346	3.44	3	3	<i>Pseudoduganella violaceinigra</i>	99.54	<i>Proteobacteria</i> <i>Betaproteobacteria</i> <i>Burkholderiales</i>
OTU.3507	2.36	1	1	<i>Bacillus spp.</i>	98.63	<i>Firmicutes</i> <i>Bacilli</i> <i>Bacillales</i>
OTU.3540	2.52	3	3	<i>Flavobacterium terrigena</i>	99.54	<i>Bacteroidetes</i> <i>Flavobacteria</i> <i>Flavobacteriales</i>
OTU.360	2.98	3	3	<i>Flavisolibacter ginsengisoli</i>	95.0	<i>Bacteroidetes</i> <i>Sphingobacteriia</i> <i>Sphingobacteriales</i>
OTU.369	5.05	1	1	<i>Paenibacillus sp. D75</i> , <i>Paenibacillus glycansilyticus</i>	100.0	<i>Firmicutes</i> <i>Bacilli</i> <i>Bacillales</i>
OTU.37	2.68	7	7	<i>Phycicola gilvus</i> , <i>Microterricola viridarii</i> , <i>Frigoribacterium faeni</i> , <i>Frondihabitans sp. RS-15</i> , <i>Frondihabitans australicus</i>	100.0	<i>Actinobacteria</i> <i>Micrococcales</i> <i>Microbacteriaceae</i>
OTU.394	4.06	1	1	<i>Paenibacillus pocheonensis</i>	100.0	<i>Firmicutes</i> <i>Bacilli</i> <i>Bacillales</i>
OTU.4	2.84	7	7, 14	<i>Agromyces ramosus</i>	100.0	<i>Actinobacteria</i> <i>Micrococcales</i> <i>Microbacteriaceae</i>
OTU.4446	3.49	7	7	<i>Catenuloplanes niger</i> , <i>Catenuloplanes castaneus</i> , <i>Catenuloplanes atrovinosus</i> , <i>Catenuloplanes crispus</i> , <i>Catenuloplanes nepalensis</i> , <i>Catenuloplanes japonicus</i>	97.72	<i>Actinobacteria</i> <i>Frankiales</i> <i>Nakamurellaceae</i>
OTU.4743	2.24	1	1	<i>Lysinibacillus fusiformis</i> , <i>Lysinibacillus sphaericus</i>	99.09	<i>Firmicutes</i> <i>Bacilli</i> <i>Bacillales</i>
OTU.48	2.99	1	1, 3	<i>Aeromonas spp.</i>	100.0	<i>Proteobacteria</i> <i>Gammaproteobacteria</i> <i>aaa34a10</i>
OTU.5	3.69	7	7	<i>Delftia tsuruhatensis</i> , <i>Delftia lacustris</i>	100.0	<i>Proteobacteria</i> <i>Betaproteobacteria</i> <i>Burkholderiales</i>
OTU.5284	3.56	7	7	<i>Isoptericola nanjingensis</i> , <i>Isoptericola hypogaeus</i> , <i>Isoptericola variabilis</i>	98.63	<i>Actinobacteria</i> <i>Micrococcales</i> <i>Promicromonosporaceae</i>
OTU.5603	3.96	1	1	<i>Paenibacillus uliginis</i>	100.0	<i>Firmicutes</i> <i>Bacilli</i> <i>Bacillales</i>
OTU.57	4.39	1	1, 3, 7, 14, 30	<i>Paenibacillus castaneae</i>	98.62	<i>Firmicutes</i> <i>Bacilli</i> <i>Bacillales</i>
OTU.5906	3.16	3	3	<i>Terrimonas sp. M-8</i>	96.8	<i>Bacteroidetes</i> <i>Sphingobacteriia</i> <i>Sphingobacteriales</i>
OTU.6	3.24	3	3	<i>Cellvibrio fulvus</i>	100.0	<i>Proteobacteria</i> <i>Gammaproteobacteria</i> <i>Pseudomonadales</i>

Table S1 – continued from previous page

OTU ID	Fold change	Day	All days	Top BLAST hits	BLAST %ID	Phylum;Class;Order
OTU.62	2.57	7	7	<i>Nakamurella flava</i>	100.0	<i>Actinobacteria Frankiales Nakamurellaceae</i>
OTU.6203	3.32	3	3	<i>Flavobacterium granuli</i> , <i>Flavobacterium glaciei</i>	100.0	<i>Bacteroidetes Flavobacteria Flavobacteriales</i>
OTU.68	3.74	7	7	<i>Shigella flexneri</i> , <i>Escherichia fergusonii</i> , <i>Escherichia coli</i> , <i>Shigella sonnei</i>	100.0	<i>Proteobacteria Gammaproteobacteria Enterobacteriales</i>
OTU.760	2.89	3	3	<i>Dyadobacter hamtensis</i>	98.63	<i>Bacteroidetes Cytophagia Cytophagales</i>
OTU.8	2.26	1	1	<i>Bacillus niaci</i>	100.0	<i>Firmicutes Bacilli Bacillales</i>
OTU.843	3.62	1	1	<i>Paenibacillus agaricidens</i>	100.0	<i>Firmicutes Bacilli Bacillales</i>
OTU.9	2.04	1	1	<i>Bacillus megaterium</i> , <i>Bacillus flexus</i>	100.0	<i>Firmicutes Bacilli Bacillales</i>

<sup>a</sup> Maximum observed  $\log_2$  of fold change.<sup>b</sup> Day of maximum fold change.<sup>c</sup> All response days.

Table S2: <sup>13</sup>C-cellulose responders BLAST against Living Tree Project

OTU ID	Fold change <sup>a</sup>	Day <sup>b</sup>	All days <sup>c</sup>	Top BLAST hits	BLAST %ID	Phylum;Class;Order
OTU.100	2.66	14	14	<i>Pseudoxanthomonas sacheonensis</i> , <i>Pseudoxanthomonas dokdonensis</i>	100.0	Proteobacteria Gammaproteobacteria Xanthomonadales
OTU.1023	4.61	30	30	No hits of at least 90% identity	80.54	Verrucomicrobia Spartobacteria Chthoniobacterales
OTU.1065	5.31	14	14, 30	No hits of at least 90% identity	84.55	Planctomycetes Planctomycetacia Planctomycetales
OTU.1087	4.32	14	14, 30	<i>Devosia soli</i> , <i>Devosia crocina</i> , <i>Devosia riboflavina</i>	99.09	Proteobacteria Alphaproteobacteria Rhizobiales
OTU.1094	3.69	30	30	<i>Sporocytophaga myxococcoides</i>	99.55	Bacteroidetes Cytophagia Cytophagales
OTU.11	3.41	14	14	<i>Stenotrophomonas pavani</i> , <i>Stenotrophomonas maltophilia</i> , <i>Pseudomonas geniculata</i>	99.54	Proteobacteria Gammaproteobacteria Xanthomonadales
OTU.114	2.78	14	14	<i>Herbaspirillum sp. SUEMI03</i> , <i>Herbaspirillum sp. SUEMI10</i> , <i>Oxalicibacterium solurbis</i> , <i>Herminiumonas fonticola</i> , <i>Oxalicibacterium horti</i>	100.0	Proteobacteria Betaproteobacteria Burkholderiales
OTU.119	3.31	14	14, 30	<i>Brevundimonas alba</i>	100.0	Proteobacteria Alphaproteobacteria Caulobacterales
OTU.120	4.76	14	14, 30	<i>Vampirovibrio chlorellavorus</i>	94.52	Cyanobacteria SM1D11 uncultured-bacterium
OTU.1204	4.32	30	30	<i>Planctomyces limnophilus</i>	91.78	Planctomycetes Planctomycetacia Planctomycetales
OTU.1312	4.07	30	30	<i>Paucimonas lemoignei</i>	99.54	Proteobacteria Betaproteobacteria Burkholderiales
OTU.132	2.81	14	14	<i>Streptomyces spp.</i>	100.0	Actinobacteria Streptomycetales Streptomycetaceae
OTU.150	4.06	14	14	No hits of at least 90% identity	86.76	Planctomycetes Planctomycetacia Planctomycetales
OTU.1533	3.43	30	30	No hits of at least 90% identity	82.27	Verrucomicrobia Spartobacteria Chthoniobacterales
OTU.154	3.24	14	14	<i>Pseudoxanthomonas mexicana</i> , <i>Pseudoxanthomonas japonensis</i>	100.0	Proteobacteria Gammaproteobacteria Xanthomonadales
OTU.165	3.1	14	14	<i>Rhizobium skieniewicense</i> , <i>Rhizobium vignae</i> , <i>Rhizobium larrymoorei</i> , <i>Rhizobium alkalisolii</i> , <i>Rhizobium galegae</i> , <i>Rhizobium huautlense</i>	100.0	Proteobacteria Alphaproteobacteria Rhizobiales
OTU.1754	4.48	14	14	<i>Asticcacaulis biprosthecum</i> , <i>Asticcacaulis benevestitus</i>	96.8	Proteobacteria Alphaproteobacteria Caulobacterales
OTU.185	4.37	14	14, 30	No hits of at least 90% identity	85.14	Verrucomicrobia Spartobacteria Chthoniobacterales
OTU.19	2.44	14	14	<i>Rhizobium alamii</i> , <i>Rhizobium mesosinicum</i> , <i>Rhizobium mongolense</i> , <i>Arthrobacter viscosus</i> , <i>Rhizobium sullae</i> , <i>Rhizobium yanglingense</i> , <i>Rhizobium loessense</i>	99.54	Proteobacteria Alphaproteobacteria Rhizobiales

Table S2 – continued from previous page

OTU ID	Fold change	Day	All days	Top BLAST hits	BLAST %ID	Phylum;Class;Order
OTU.2192	3.49	30	14, 30	No hits of at least 90% identity	83.56	Verrucomicrobia Spartobacteria Chthoniobacteriales
OTU.228	2.54	30	30	<i>Sorangium cellulosum</i>	98.17	Proteobacteria Deltaproteobacteria Myxococcales
OTU.241	2.66	14	14	No hits of at least 90% identity	87.73	Verrucomicrobia Spartobacteria Chthoniobacteriales
OTU.257	2.94	14	14	<i>Lentzea waywayandensis</i> , <i>Lentzea flaviverrucosa</i>	100.0	Actinobacteria Pseudonocardiales Pseudonocardiaceae
OTU.266	4.54	30	14, 30	No hits of at least 90% identity	83.64	Verrucomicrobia Spartobacteria Chthoniobacteriales
OTU.28	2.59	14	14	<i>Rhizobium giardinii</i> , <i>Rhizobium tubonense</i> , <i>Rhizobium tibeticum</i> , <i>Rhizobium mesoamericanum CCGE 501</i> , <i>Rhizobium herbae</i> , <i>Rhizobium endophyticum</i>	99.54	Proteobacteria Alphaproteobacteria Rhizobiales
OTU.285	3.55	30	14, 30	<i>Blastopirellula marina</i>	90.87	Planctomycetes Planctomycetacia Planctomycetales
OTU.32	2.34	3	3	<i>Sandaracinus amyloyticus</i>	94.98	Proteobacteria Deltaproteobacteria Myxococcales
OTU.327	2.99	14	14	<i>Asticcacaulis biprostheciun</i> , <i>Asticcacaulis benevestitus</i>	98.63	Proteobacteria Alphaproteobacteria Caulobacteriales
OTU.351	3.54	14	14, 30	<i>Pirellula staleyi DSM 6068</i>	91.86	Planctomycetes Planctomycetacia Planctomycetales
OTU.3594	3.83	30	30	<i>Chondromyces robustus</i>	90.41	Proteobacteria Deltaproteobacteria Myxococcales
OTU.3775	3.88	14	14	<i>Devasia glacialis</i> , <i>Devasia chinhatensis</i> , <i>Devasia geoensis</i> , <i>Devasia yakushimensis</i>	98.63	Proteobacteria Alphaproteobacteria Rhizobiales
OTU.429	3.7	30	14, 30	<i>Devasia limi</i> , <i>Devasia psychrophila</i>	97.72	Proteobacteria Alphaproteobacteria Rhizobiales
OTU.4322	4.19	14	7, 14, 30	No hits of at least 90% identity	89.14	Chloroflexi Herpetosiphonales Herpetosiphonaceae
OTU.442	3.05	30	30	<i>Chondromyces robustus</i>	92.24	Proteobacteria Deltaproteobacteria Myxococcales
OTU.465	3.79	30	30	<i>Ohtaekwangia kribbensis</i>	92.73	Bacteroidetes Cytophagia Cytophagales
OTU.473	3.58	14	14	<i>Pirellula staleyi DSM 6068</i>	90.91	Planctomycetes Planctomycetacia Planctomycetales
OTU.484	4.92	14	14, 30	No hits of at least 90% identity	89.09	Planctomycetes Planctomycetacia Planctomycetales
OTU.5	2.69	14	14	<i>Delftia tsuruhatensis</i> , <i>Delftia lacustris</i>	100.0	Proteobacteria Betaproteobacteria Burkholderiales
OTU.518	4.8	14	14	<i>Hydrogenophaga intermedia</i>	100.0	Proteobacteria Betaproteobacteria Burkholderiales
OTU.5190	3.6	30	14, 30	No hits of at least 90% identity	88.13	Chloroflexi Herpetosiphonales Herpetosiphonaceae
OTU.541	4.49	30	30	No hits of at least 90% identity	84.23	Verrucomicrobia Spartobacteria Chthoniobacteriales
OTU.5539	4.01	14	14	<i>Devasia subaequoris</i>	98.17	Proteobacteria Alphaproteobacteria Rhizobiales
OTU.573	3.03	30	30	<i>Adhaeribacter aerophilus</i>	92.76	Bacteroidetes Cytophagia Cytophagales

Table S2 – continued from previous page

OTU ID	Fold change	Day	All days	Top BLAST hits	BLAST %ID	Phylum;Class;Order
OTU.6	3.62	7	3, 7, 14	<i>Cellvibrio fulvus</i>	100.0	Proteobacteria Gammaproteobacteria Pseudomonadales
OTU.600	3.48	30	30	No hits of at least 90% identity	80.37	Planctomycetes Planctomycetacia Planctomycetales
OTU.6062	4.83	30	30	<i>Dokdonella sp. DC-3</i> , <i>Luteibacter rhizovicinus</i>	97.26	Proteobacteria Gammaproteobacteria Xanthomonadales
OTU.627	4.43	14	14	<i>Verrucomicrobiaceae bacterium DC2a-G7</i>	100.0	Verrucomicrobia Verrucomicrobiae Verrucomicrobiales
OTU.633	3.84	30	30	No hits of at least 90% identity	89.5	Proteobacteria Deltaproteobacteria Myxococcales
OTU.638	4.0	30	30	<i>Luteolibacter sp. CCTCC AB 2010415</i> , <i>Luteolibacter algae</i>	93.61	Verrucomicrobia Verrucomicrobiae Verrucomicrobiales
OTU.64	4.31	14	7, 14, 30	No hits of at least 90% identity	89.5	Chloroflexi Herpetosiphonales Herpetosiphonaceae
OTU.663	3.63	30	30	<i>Pirellula staleyi DSM 6068</i>	90.87	Planctomycetes Planctomycetacia Planctomycetales
OTU.669	3.34	30	30	<i>Ohtaekwangia koreensis</i>	92.69	Bacteroidetes Cytophagia Cytophagales
OTU.670	2.87	30	30	<i>Adhaeribacter aerophilus</i>	91.78	Bacteroidetes Cytophagia Cytophagales
OTU.766	3.21	14	14, 30	<i>Devosia insulae</i>	99.54	Proteobacteria Alphaproteobacteria Rhizobiales
OTU.83	5.61	14	7, 14, 30	<i>Luteolibacter sp. CCTCC AB 2010415</i>	97.72	Verrucomicrobia Verrucomicrobiae Verrucomicrobiales
OTU.862	5.87	14	14	<i>Allakutzneria albata</i>	100.0	Actinobacteria Pseudonocardiales Pseudonocardiaceae
OTU.899	2.28	30	30	<i>Enhygromyxa salina</i>	97.72	Proteobacteria Deltaproteobacteria Myxococcales
OTU.90	2.94	14	14, 30	<i>Sphingopyxis panaciterra</i> , <i>Sphingopyxis chilensis</i> , <i>Sphingopyxis sp. BZ30</i> , <i>Sphingomonas sp.</i>	100.0	Proteobacteria Alphaproteobacteria Sphingomonadales
OTU.900	4.87	14	14	<i>Brevundimonas vesicularis</i> , <i>Brevundimonas nasdae</i>	100.0	Proteobacteria Alphaproteobacteria Caulobacterales
OTU.971	3.68	30	30	No hits of at least 90% identity	78.57	Chloroflexi Anaerolineae Anaerolineales
OTU.98	3.68	14	7, 14, 30	No hits of at least 90% identity	88.18	Chloroflexi Herpetosiphonales Herpetosiphonaceae
OTU.982	4.47	14	14	<i>Devosia neptuniae</i>	100.0	Proteobacteria Alphaproteobacteria Rhizobiales

<sup>a</sup> Maximum observed  $\log_2$  of fold change.<sup>b</sup> Day of maximum fold change.<sup>c</sup> All response days.

1222

# Supplemental Information

1223

## 1224 Contents

1225	<b>1 Supplemental Methods</b>	<b>1</b>
1226	1.1 Soil Collection and Preparation . . . . .	1
1227	1.2 Cellulose production . . . . .	2
1228	1.3 Soil microcosms . . . . .	2
1229	1.4 Nucleic acid extraction . . . . .	3
1230	1.5 Isopycnic centrifugation and fractionation . . . . .	3
1231	1.6 DNA Sequencing . . . . .	4
1232	1.6.1 PCR amplification of SSU rRNA genes . . . . .	4
1233	1.6.2 DNA sequence quality control . . . . .	4
1234	1.6.3 OTU binning . . . . .	5
1235	1.6.4 Phylogenetic reconstruction . . . . .	5
1236	1.6.5 Ordination and statistical analysis of differences in SSU rRNA gene composition	5
1237	1.7 OTU characteristics . . . . .	5
1238	1.7.1 Identifying <sup>13</sup> C responders . . . . .	5
1239	1.7.2 Estimating rrn copy number . . . . .	6
1240	1.7.3 NRI, NTI, and consenTRAIT . . . . .	6
1241	1.7.4 Buoyant density shift estimates . . . . .	6
1242	1.7.5 Finding cultured relatives of OTUs . . . . .	6
1243	1.7.6 OTU changes in relative abundance with time . . . . .	7
1244	1.8 Sequencing and density fractionation statistics . . . . .	7

## 1245 1 Supplemental Methods

### 1246 1.1 Soil Collection and Preparation

1247 We collected soils from an organic farm in Penn Yan, New York. Soils were Honoeye/Lima, a silty  
 1248 clay loam on calcareous bedrock. To get a field average, cores (5 cm diameter x 10 cm depth) were  
 1249 collected in duplicate from six different sampling locations around the field using a slide hammer  
 1250 bulk density sampler (coordinates: (1) N 42° 40.288 W 77° 02.438, (2) N 42 40.296 W 77° 02.438,  
 1251 (3) N 42° 40.309 W 77° 02.445, (4) N 42° 40.333 W 77° 02.425, (5) N 42° 40.340 W 77° 02.420, (6)  
 1252 N 42° 40.353 W 77° 02.417) on November 21, 2011. Soil cores were sieved (2mm), homogenized by  
 1253 mixing, and stored at 4 °C until pre-incubation (within 1-2 week of collection). Carbon (C) and  
 1254 nitrogen (N) content were previously measured for these soils [1]. Reported soil C values for the  
 1255 organic field were 12.15 ( $\pm$  s.d. 0.78) mg C g<sup>-1</sup> dry soil and 1.16 ( $\pm$  s.d. 0.13) mg N g<sup>-1</sup> dry soil  
 1256 [1].

1257 **1.2 Cellulose production**

1258 Bacterial cellulose was produced by *Gluconoacetobacter xylinus* grown in Heo and Son [2] minimal  
1259 media (HS medium) made with 0.1% glucose and without inositol. For the production of  
1260  $^{13}\text{C}$ -cellulose,  $^{13}\text{C}_6\text{-D-glucose}$ , 99 atom %  $^{13}\text{C}$  (Isotec) was used. Cellulose was produced in 1L  
1261 Erlenmeyer flasks containing 100 mL HS medium inoculated with three colonies of *Gluconoacetobacter*  
1262 *xylinus* grown on HS agar plates. Flasks were incubated statically in the dark at 30°C for 2-3  
1263 weeks. Cellulose pellicles were decanted, rinsed with deionized water, suspended in two volumes  
1264 of 1% Alconox, and then autoclaved. Cellulose pellicles were purified by dialysis for 12 hr in 1  
1265 L deionized water and dialysis was repeated 10 times. Harvested pellicles were dried overnight  
1266 (60°C), cut into pieces, and ground using a 5100 Mixer/Mill (SPEX SamplePrep, Metuchen, NJ),  
1267 and dry sieved to 53  $\mu\text{m}$ -250  $\mu\text{m}$ . The particulate size range was selected to be representative of  
1268 particulate organic matter in soils (3).

1269 The purity of ground cellulose was checked by biological assay, Benedict's reducing sugars assay,  
1270 Bradford assay, and isotopic analysis. *E. coli* is not able to use cellulose as a C source but is capable  
1271 of growth on a variety of nutrients available in the HS medium. The biological assay consisted of *E.*  
1272 *coli* inoculated into minimal M9 media which lacked a C source and was supplemented with either:  
1273 (1) 0.01% glucose, (2) 2.5 mg purified, ground cellulose, (3) 25 mg purified, ground cellulose, (4)  
1274 25 mg purified, ground cellulose and 0.01% glucose. Growth in media was checked by spectrometer  
1275 ( $\text{OD}_{450}$ ). No measurable growth was observed with either 2 mg or 25 mg cellulose, indicating  
1276 absence of contaminating nutrients that can support growth of *E. coli*. In addition, the presence  
1277 of 25 mg cellulose did not inhibit the growth of *E. coli* cultures provided with glucose (relative to  
1278 control), indicating the absence of compounds in the purified cellulose that could inhibit microbial  
1279 growth.

1280 Purified cellulose was also assayed for residual proteins and sugars using Bradford and Benedict's  
1281 assays, respectively. Bradford assay was performed as in [3]. Ground, purified cellulose contained  
1282 6.92  $\mu\text{g}$  protein mg cellulose $^{-1}$  (*i.e.* 99.31% purity). Reducing sugars were not detected in cellulose  
1283 using Benedict's reducing sugar assay [4] tested at 10 mg cellulose ml $^{-1}$ . Finally,  $^{13}\text{C}$ -cellulose had  
1284 an average  $96\% \pm 5$  (s.d.) degree of  $^{13}\text{C}$  labeling as determined by isotopic analysis (UCDavis  
1285 Stable Isotope Facility).

1286 **1.3 Soil microcosms**

1287 Microcosms were created by adding 10 g d.w. sieved soil to a 250 mL Erlenmeyer flask capped  
1288 with a butyl rubber stopper. The headspace was flushed with air every 3 days which was sufficient  
1289 to prevent anoxia (data not shown). Microcosms were pre-incubated at room temperature for  
1290 2 weeks until the soil respiration rate (determined by GCMS measurement of head space CO $_2$ )  
1291 had stabilized. Sieving causes a transient increase in soil respiration rate presumably due to the  
1292 liberation of fresh labile soil organic matter [5]. Pre-incubation ensures that this labile organic  
1293 matter is consumed and/or stabilized prior to the beginning of the experiment. Respiration rate  
1294 (CO $_2$ ) stabilized after 10 days (data not shown).

1295 Three parallel treatments were established. Each treatment received the same amendment,  
1296 where the only difference was the isotopically labeled component in the amendment. The treat-  
1297 ments included an unlabeled control treatment and treatments that substituted either  $^{13}\text{C}$ -cellulose  
1298 (synthesized as described above) or  $^{13}\text{C}_5\text{-D-xylose}$  (98 atom %  $^{13}\text{C}$  (Isotec)) for their unlabeled  
1299 equivalents. The molecular composition of the amendment was designed to approximate switch-  
1300 grass biomass with hemicellulose replaced by its constituent monomers [6, 7]. The amendment was

1301 added at 5.3 mg g<sup>-1</sup> d.w. soil which is representative of natural concentrations in soil during early  
1302 phases of decomposition [8]. The amendment contained by mass: 38% cellulose, 23% lignin, 20%  
1303 xylose, 3% arabinose, 1% galactose, 1% glucose, and 0.5% mannose, 10.6% amino acids (Teknova  
1304 C0705), and 2.9% Murashige Skoog basal salt mixture which contains macro- and micro-nutrients  
1305 that are associated with plant biomass (Sigma Aldrich M5524). The amendment had a C:N ratio of  
1306 10. Cellulose (2 mg cellulose g<sup>-1</sup> d.w. soil) and lignin (1.2 mg lignin g<sup>-1</sup> d.w. soil) were uniformly  
1307 distributed over the soil surface as a powder and the remaining constituents were added in solution  
1308 in a volume of 0.12 ml g<sup>-1</sup> d.w. soil. The volume of liquid was determined in relation to soil  
1309 moisture to achieve 50% water holding capacity. Water holding capacity of 50% was chosen, in re-  
1310 lation to the texture for this soil, to achieve approximately 70% water filled pore space, which is the  
1311 optimal water content for respiration [9]. A total of 12 microcosms were established per treatment.  
1312 Microcosms were sampled destructively on days 1, 3, 7, 14, and 30 and soils were frozen at -80  
1313 °C. The cellulose treatment was not sampled on day 1 because it was not expected that significant  
1314 cellulose metabolism would have occurred within this time. The abbreviation 13CXPS refers to the  
1315 <sup>13</sup>C-xylose treatment (<sup>13</sup>C Xylose Plant Simulant), 13CCPS refers to the <sup>13</sup>C-cellulose treatment  
1316 and 12CCPS refers to the unlabeled control. A subset of soil from each sample was reserved for  
1317 isotopic analysis at the Cornell University Stable Isotope Laboratory to determine the mass of <sup>13</sup>C  
1318 remaining in soil.

#### 1319 1.4 Nucleic acid extraction

1320 Nucleic acids were extracted from 0.25 g soil using a modified Griffiths protocol [10]. Cells were  
1321 lysed by bead beating for 1 min at 5.5 m s<sup>-1</sup> in 2 mL lysis tubes containing 0.5 g of 0.1 mm diameter  
1322 silica/zirconia beads (treated at 300 °C for 4 hours to remove RNases), 0.5 mL extraction buffer  
1323 (240 mM Phosphate buffer 0.5% N-lauryl sarcosine), and 0.5 mL phenol-chloroform-isoamyl alcohol  
1324 (25:24:1) for 1 min at 5.5 m s<sup>-1</sup>. After lysis, 85 μL 5 M NaCl and 60 μL 10% hexadecyltrimethylammonium  
1325 bromide (CTAB)/0.7 M NaCl were added to lysis tube, vortexed, chilled for 1 min on ice, and  
1326 centrifuged at 16,000 x g for 5 min at 4C. The aqueous layer was transferred to a new tube and  
1327 reserved on ice. To increase DNA recovery, the pellet was back extracted with 85 μL 5 M NaCl and  
1328 0.5 mL extraction buffer. The aqueous extract was washed with 0.5 mL chloroform:isoamyl alcohol  
1329 (24:1). Nucleic acids were precipitated by addition of 2 volumes polyethylene glycol solution (30%  
1330 PEG 8000, 1.6 M NaCl) on ice for 2 hrs, followed by centrifugation at 16,000 x g, 4C for 30 min.  
1331 The supernatant was discarded and pellets were washed with 1 mL ice cold 70% EtOH. Pellets  
1332 were air dried, resuspended in 50 μL TE and stored at -20 °C. To prepare nucleic acid extracts  
1333 for isopycnic centrifugation as previously described [11], DNA was size selected (> 4kb) using 1%  
1334 low melt agarose gel and β-agarase I enzyme extraction per manufacturers protocol (New England  
1335 Biolab, M0392S). Final resuspension of DNA pellet was in 50 μL TE.

#### 1336 1.5 Isopycnic centrifugation and fractionation

1337 We fractionated DNA on density gradients for <sup>13</sup>C-xylose treatments (days 1, 3, 7, 14, 30), <sup>13</sup>C-  
1338 cellulose treatments (days 3, 7, 14, 30), and control treatments (days 1, 3, 7, 14, 30). A total of  
1339 5 μg DNA was added to each 4.7 mL CsCl density gradient. Density gradient were composed of  
1340 1.69 g mL<sup>-1</sup> CsCl ml<sup>-1</sup> in gradient buffer solution (pH 8.0 15 mM Tris-HCl, 15 mM EDTA, 15  
1341 mM KCl). Centrifugation was performed at 55,000 rpm 20 °C for 66 hr using a TLA-110 rotor in a  
1342 Beckman Coulter Optima MAX-E ultracentrifuge. Fractions of ~100 μL were collected from below

1343 by displacing the DNA-CsCl-gradient buffer solution in the centrifugation tube with water using a  
1344 syringe pump at a flow rate of  $3.3 \mu\text{L s}^{-1}$  [12]. Fractions were collected in Acroprep 96 filter plates  
1345 (part no. 5035, Pall Life Sciences). The refractive index of each fraction was measured using a  
1346 Reichart AR200 digital refractometer modified as previously described to measure a volume of  $5 \mu\text{L}$   
1347 [11]. Buoyant density was calculated from the refractive index as previously described [11] using the  
1348 equation  $\rho = a\eta - b$ , where  $\rho$  is the density of the CsCl ( $\text{g mL}^{-1}$ ),  $\eta$  is the measured refractive index,  
1349 and  $a$  and  $b$  are coefficient values of 10.9276 and 13.593, respectively, for CsCl at  $20^\circ\text{C}$  [13] and  
1350 correcting for non-CsCl salts in the gradient buffer. A total of 35 fractions were collected from each  
1351 gradient and the average density between fractions was  $0.0040 \text{ g mL}^{-1}$ . The DNA was desalted by  
1352 washing with TE (5X 200  $\mu\text{L}$ ) in the Acroprep filter wells. DNA was resuspended in 50  $\mu\text{L}$  TE.

## 1353 1.6 DNA Sequencing

### 1354 1.6.1 PCR amplification of SSU rRNA genes

1355 SSU rRNA genes were amplified from gradient fractions ( $n = 20$  per gradient) and from non-  
1356 fractionated DNA from soil. Barcoded primers consisted of: 454-specific adapter B, a 10 bp barcode  
1357 (Reference 90), a 2 bp linker (5-CA-3), and 806R primer for reverse primer (BA806R); and 454-  
1358 specific adapter A, a 2 bp linker (5-TC-3), and 515F primer for forward primer (BA515F). Each  
1359 PCR contained 1.25 U l-1 AmpliTaq Gold (Life Technologies, Grand Island, NY; N8080243), 1X  
1360 Buffer II (100 mM Tris-HCl, 500 mM KCl, pH 8.3), 2.5 mM MgCl<sub>2</sub>, 200 M of each dNTP, 0.5 mg  
1361 ml-1 BSA, 0.2 M BA515F, 0.2 M BA806R, and 10 L of 1:30 DNA template in 25 l total volume).  
1362 The PCR conditions were 95 °C for 5min followed by 22 cycles of 95°C for 10 s, 53 °C for 30 s,  
1363 and 72 °C for 30 s, followed by a final elongation at 72 °C for 5 min. Amplification products were  
1364 checked by 1% agarose gel. Reactions were performed in triplicate and pooled. Amplified DNA was  
1365 gel purified (1% low melt agarose) using the Wizard SV gel and PCR clean-up system (Promega,  
1366 Madison, WI; A9281) per manufacturers protocol. Samples were normalized by SeqalPrepTM  
1367 normalization plates (Invitrogen, Carlsbad, CA; A10510) and pooled in equimolar concentration.  
1368 Amplicons were sequenced on Roche 454 FLX system using titanium chemistry at Selah Genomics  
1369 (Columbia, SC).

### 1370 1.6.2 DNA sequence quality control

1371 SSU rRNA gene sequences were initially screened by maximum expected errors at a specific read  
1372 length threshold [14]. Reads that had more than 0.5 expected errors at a length of 250 nt were  
1373 discarded. The remaining reads were aligned to the Silva Reference Alignment as provided in the  
1374 Mothur software package using the Mothur NAST aligner [15, 16]. Reads that did not align to  
1375 the expected region of the SSU rRNA gene were discarded. After expected error and alignment  
1376 based quality control. The remaining quality controlled reads were annotated using the UClust  
1377 taxonomic annotation framework in [17, 18]. We used 97% cluster seeds from the Silva SSU rRNA  
1378 database (release 111Ref) [19] as reference for taxonomic annotation (provided on the QIIME  
1379 website) [19]. Quality control screening filtered out 344,472 of 1,720,480 raw sequencing reads  
1380 leaving 1,376,008 reads for downstream analyses. Reads annotated as "Chloroplast", "Eukaryota",  
1381 "Archaea", "Unassigned" or "mitochondria" were culled from the dataset.

1382 **1.6.3 OTU binning**

1383 Sequences were distributed into OTUs with a centroid based clustering algorithm (i.e. UPARSE  
1384 [14]). The centroid selection also included robust chimera screening [14]. OTU centroids were  
1385 established at a threshold of 97% sequence identity and non-centroid sequences were mapped back  
1386 to centroids. Reads that could not be mapped to an OTU centroid at greater than or equal to 97%  
1387 sequence identity were discarded.

1388 **1.6.4 Phylogenetic reconstruction**

1389 We used SSU-Align [20, 21] to align SSU rRNA gene sequences. Columns in the alignment that  
1390 were aligned with poor confidence (< 95% of characters had posterior probability > 95%) were  
1391 not considered when building the phylogenetic tree leaving a multiple sequence alignment of 216  
1392 columns. Additionally, the alignment was trimmed to coordinates such that all sequences in the  
1393 alignment began and ended at the same positions. FastTree [22] was used with default parameters  
1394 to build the phylogeny.

1395 **1.6.5 Ordination and statistical analysis of differences in SSU rRNA gene composition**

1396 NMDS ordination was performed on weighted Unifrac [23] distances between samples. The Phyloseq  
1397 [24] wrapper for Vegan [25] (both R packages) was used to compute sample values along NMDS  
1398 axes. The 'adonis' function in Vegan was used to perform Adonis tests (default parameters) [26].

1399 **1.7 OTU characteristics**

1400 **1.7.1 Identifying  $^{13}\text{C}$  responders**

1401 Figures S11 and S12 demonstrate raw data for responder and non-responder OTUs, respectively.  
1402 Responders increased in relative abundance in the high density fractions due to  $^{13}\text{C}$ -labeling of  
1403 their DNA. As our data is compositional, often OTUs had consistent *relative* abundance across  
1404 the density gradients indicating the OTU DNA concentration across the gradient mirrored that  
1405 of the total DNA concentration. If OTU DNA is centered outside the main distribution of DNA  
1406 due to G+C content and/or  $^{13}\text{C}$ -labeling its relative abundance increases near the center of the  
1407 OTU DNA concentration profile. Thus, we identified responders by finding OTUs enriched in high  
1408 density fractions of  $^{13}\text{C}$ -treatment gradients relative to control. This technique accounts for the  
1409 variation in OTU base abundance and the variation in OTU G+C content (and therefore natural  
1410 buoyant density) because relative abundances in gradient fractions from  $^{13}\text{C}$ -treatments are always  
1411 compared to those in corresponding gradient fractions from control gradients.

1412 We used DESeq2 (R package), an RNA-Seq differential expression statistical framework [27], to  
1413 identify OTUs that were enriched in high density gradient fractions from  $^{13}\text{C}$ -treatments relative  
1414 to corresponding gradient fractions from control treatments (for review of RNA-Seq differential  
1415 expression statistics applied to microbiome OTU count data see [28]). We define "high density  
1416 gradient fractions" as gradient fractions whose density falls between 1.7125 and 1.755 g ml<sup>-1</sup>.  
1417 Briefly, DESeq2 includes several features that enable robust estimates of standard error in addition  
1418 to reliable ranking of logarithmic fold change (LFC) (i.e. gamma-Poisson regression coefficients)  
1419 in OTU relative abundance even with low count OTUs where LFC can often be noisy. Further,  
1420 statistical evaluation of LFC can be performed with user-selected thresholds, as opposed to the  
1421 typical null hypothesis that LFC is exactly zero, enabling the most biologically interesting OTUs

1422 to be identified for subsequent analyses. For each OTU, we calculated LFC and corresponding  
1423 standard errors for enrichment in high density gradient fractions of  $^{13}\text{C}$  treatments relative to  
1424 control. Subsequently, a one-sided Wald test was used to statistically assess LFC values. The user-  
1425 defined null hypothesis was that LFC was less than one standard deviation above the mean of all  
1426 LFC values. P-values were corrected for multiple comparisons using the Benjamini and Hochberg  
1427 method [29]. We independently filtered OTUs on the basis of sparsity prior to correcting P-values  
1428 for multiple comparisons. The sparsity value that yielded the most adjusted P-values less than 0.10  
1429 was selected for independent filtering by sparsity. Briefly, OTUs were eliminated if they failed to  
1430 appear in at least 45% of high density gradient fractions for a given  $^{13}\text{C}$ /control treatment pair.  
1431 These sparse OTUs are unlikely to have sufficient data to allow for the determination of statistical  
1432 significance. We selected a false discovery rate of 10% to denote statistical significance.

#### 1433 1.7.2 Estimating *rrn* copy number

1434 We estimated the *rrn* copy number for each OTU as described [30] (i.e. we used the code and  
1435 reference information provided by the authors [30] directly). In brief, OTU centroid sequences were  
1436 inserted into a reference SSU rRNA gene phylogeny [31] from organisms of known *rrn* copy number.  
1437 The *rrn* copy number was then inferred from the phylogenetic placement in the reference phylogeny.

#### 1438 1.7.3 NRI, NTI, and consenTRAIT

1439 NRI and NTI were calculated using the “picante” R package [32]. We used the “independentswap”  
1440 null model for phylogenetic distribution. The consenTRAIT clade depth for xylose and cellulose  
1441 responders was calculated using R code from the original publication describing the metric [33]  
1442 which employs the R “adephylo” package [34].

#### 1443 1.7.4 Buoyant density shift estimates

1444 DNA buoyant density (BD) increases with atom %  $^{13}\text{C}$ . Therefore, the magnitude of  $\Delta\hat{BD}$  indicates  
1445 the degreee of isotopic labeling for an OTU. We measured  $\Delta\hat{BD}$  as the change in an OTU’s density  
1446 profile center of mass between corresponding control and labeled gradients (Figure S11). Because  
1447 all gradients did not span the same density range and gradient fractions cannot be taken at specific  
1448 density positions, we limited our  $\Delta\hat{BD}$  analysis to the density range for where all density gradients  
1449 overlapped. Within this density range we linearly interpolated 20 evenly spaced relative abundance  
1450 values. The center of mass for an OTU along the density gradient was then the density weighted  
1451 average where weights were the linearly interpolated relative abundance values.  $\Delta\hat{BD}$  values are  
1452 based on relative abundance profiles and would be distorted in comparison to  $\Delta\hat{BD}$  based on  
1453 absolute DNA concentration profiles and should be interpreted with this transformation in mind.

#### 1454 1.7.5 Finding cultured relatives of OTUs

1455 OTU centroids were compared (BLAST [35, 36]) to sequences in “The All-Species Living Tree”  
1456 project (LTP). The LTP is a collection of SSU rRNA gene sequences for classified species of Archaea  
1457 and Bacteria [37]. We used LTP version 115 for analyses in this paper.

1458 **1.7.6 OTU changes in relative abundance with time**

1459 We identified OTUs that changed in relative abundance over time using DESeq2 [27]. Specifically,  
1460 we used day treated as an ordered factor as the regressor with LFC of the relative abundance in  
1461 non-fractionated DNA as the outcome in the general linear model. We used the default DESeq2  
1462 base mean independent filtering and disabled the Cook's cutoff outlier detection. The null model  
1463 was that abundance did not change with time and we assessed significance at a false discovery rate  
1464 of 10%.

1465 **1.8 Sequencing and density fractionation statistics**

1466 Microcosm DNA was density fractionated on CsCl density gradients. We sequenced SSU rRNA  
1467 gene amplicons from a total of 277 CsCl gradient fractions from 14 CsCl gradients and 12 bulk  
1468 microcosm DNA samples. The SSU rRNA gene data set contained 1,102,685 total sequences. The  
1469 average number of sequences per sample was 3,816 (sd 3,629) and 265 samples had over 1,000  
1470 sequences. We sequenced SSU rRNA gene amplicons from an average of 19.8 fractions per CsCl  
1471 gradient (sd 0.57). The average density between fractions was 0.0040 g mL<sup>-1</sup>. The sequencing effort  
1472 recovered a total of 5,940 OTUs. 2,943 of the total 5,940 OTUs were observed in bulk samples. We  
1473 observed 33 unique phylum and 340 unique genus annotations.

1474 **References**

- [1] Berthrong S-T, Buckley D-H, Drinkwater L-E (2013) Agricultural management and labile carbon additions affect soil microbial community structure and interact with carbon and nitrogen cycling. *Microb Ecol* 66(1): 158–170.
- [2] Heo M-S, Son H-J (2002) Development of an optimized, simple chemically defined medium for bacterial cellulose production by *Acetobacter* sp. A9 in shaking cultures. *Biotechnol Appl Biochem* 36(1): 41.
- [3] Bradford M-M (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 72(1-2): 248–254.
- [4] Benedict S-R (1909) A reagent for the detection of reducing sugars. *Journal of Biological Chemistry* 5(5): 485–487.
- [5] Datta R, Vranová V, Pavelka M, Rejšek K, Formánek P (2014) Effect of soil sieving on respiration induced by low-molecular-weight substrates. *International Agrophysics* 28(1): 119–124.
- [6] Yan J, Hu Z, Pu Y, Brummer E-C, Ragauskas A-J (2010) Chemical compositions of four switchgrass populations. *Biomass and Bioenergy* 34(1): 48–53.
- [7] David K, Ragauskas A-J (2010) Switchgrass as an energy crop for biofuel production: A review of its ligno-cellulosic chemical properties. *Energy Environ Sci* 3(9): 1182.
- [8] Schneckenberger K, Demin D, Stahr K, Kuzyakov Y (2008) Microbial utilization and mineralization of <sup>14</sup>C glucose added in six orders of concentration to soil. *Soil Biol Biochem* 40(8): 1981–1988.

- 1495 [9] Linn D-M, Doran J-W (1984) Aerobic and anaerobic microbial populations in no-till and  
1496 plowed soils. *Soil Sci Soc Am J* 48(4): 794.
- 1497 [10] Griffiths R-I, Whiteley A-S, O'Donnell A-G, Bailey M-J (2000) Rapid method for coextraction  
1498 of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based  
1499 microbial community composition. *Appl Environ Microbiol* 66(12): 5488–5491.
- 1500 [11] Buckley D-H, Huangyutitham V, Hsu S-F, Nelson T-A (2007) Stable isotope probing with  $^{15}\text{N}$   
1501 achieved by disentangling the effects of genome G+C content and isotope enrichment on DNA  
1502 density. *Appl Environ Microbiol* 73(10): 3189–3195.
- 1503 [12] Manefield M, Whiteley A-S, Griffiths R-I, Bailey M-J (2002) RNA Stable isotope probing a  
1504 novel means of linking microbial community function to phylogeny. *Appl Environ Microbiol*  
1505 68(11): 5367–5373.
- 1506 [13] Birnie G-D (1978) Centrifugal separations in Molecular and cell biology. (Butterworth & Co  
1507 Publishers Ltd, Boston).
- 1508 [14] Edgar R-C (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads.  
1509 *Nat. Methods* 10(10): 996–998.
- 1510 [15] DeSantis T-Z, Hugenholtz P, Keller K, Brodie E-L, Larsen N, Piceno Y-M, et al. (2006) NAST:  
1511 a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic Acids  
1512 Res* 34(suppl 2): W394–W399.
- 1513 [16] Schloss P-D, Westcott S-L, Ryabin T, Hall J-R, Hartmann M, Hollister E-B, et al. (2009)  
1514 Introducing mothur: open-source, platform-independent, community-supported software for  
1515 describing and comparing microbial communities. *Appl Environ Microbiol* 75(23): 7537–7541.
- 1516 [17] Caporaso J-G, Kuczynski J, Stombaugh J, Bittinger K, Bushman F-D, Costello E-K, et al.  
1517 (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods*  
1518 7(5): 335–336.
- 1519 [18] Edgar R-C (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*  
1520 26(19): 2460–2461.
- 1521 [19] Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. (2013) The SILVA ribosomal  
1522 RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*  
1523 41: D590–596.
- 1524 [20] Nawrocki E-P, Kolbe D-L, Eddy S-R (2009) Infernal 1.0: inference of RNA alignments. *Bioinformatics*  
1525 25(10): 1335–1337.
- 1526 [21] Nawrocki E-P, Eddy S-R (2013) Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*  
1527 29(22): 2933–2935.
- 1528 [22] Price M-N, Dehal P-S, Arkin A-P (2010) FastTree2 - approximately maximum-likelihood trees  
1529 for large alignments. *PLoS ONE* 5(3): e9490.
- 1530 [23] Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial  
1531 communities. *Appl Environ Microbiol* 71(12): 8228–8235.

- 1532 [24] McMurdie P-J, Holmes S (2013) phyloseq: an R package for reproducible interactive analysis  
1533 and graphics of microbiome census data. *PLoS ONE* 8(4): e61217.
- 1534 [25] Oksanen J, Blanchet F-G, Kindt R, Legendre P, Minchin P-R, O'Hara R-B, et al. (2015) vegan:  
1535 Community Ecology Package.
- 1536 [26] Anderson M-J (2001) A new method for non-parametric multivariate analysis of variance.  
1537 *Austral Ecol* 26(1): 32–46.
- 1538 [27] Love M-I, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for  
1539 RNA-seq data with DESeq2. *Genome Biol* 15(12): 550.
- 1540 [28] McMurdie P-J, Holmes S (2014) Waste not, want not: why rarefying microbiome data is  
1541 inadmissible. *PLoS Comput Biol* 10(4): e1003531.
- 1542 [29] Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful  
1543 approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*  
1544 57(1): 289–300.
- 1545 [30] Kembel S-W, Wu M, Eisen J-A, Green J-L (2012) Incorporating 16S gene copy number in-  
1546 formation improves estimates of microbial diversity and abundance. *PLoS Comput Biol* 8(10):  
1547 e1002743.
- 1548 [31] Matsen F-A, Kodner R-B, Armbrust E-V (2010) pplacer: linear time maximum-likelihood and  
1549 Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics*  
1550 11: 538.
- 1551 [32] Kembel S, Cowan P, Helmus M, Cornwell W, Morlon H, Ackerly D, et al. (2010) Picante: R  
1552 tools for integrating phylogenies and ecology. *Bioinformatics* 26: 1463–1464.
- 1553 [33] Martiny A-C, Treseder K, Pusch G (2013) Phylogenetic conservatism of functional traits in  
1554 microorganisms. *ISME J* 7(4): 830–838.
- 1555 [34] Jombart T, Dray S (2010) adephylo: exploratory analyses for the phylogenetic comparative  
1556 method.. *Bioinformatics* 26: 1907–1909.
- 1557 [35] Altschul S-F, Gish W, Miller W, Myers E-W, Lipman D-J (1990) Basic local alignment search  
1558 tool. *J Mol Biol* 215(3): 403–410.
- 1559 [36] Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. (2009) BLAST+:  
1560 architecture and applications. *BMC Bioinformatics* 10: 421.
- 1561 [37] Yarza P, Richter M, Peplies J, Euzeby J, Amann R, Schleifer K-H, et al. (2008) The All-Species  
1562 Living Tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst  
1563 Appl Microbiol* 31(4): 241–250.