October 2019

CHANGES IN SOIL MICROBIAL COMMUNITIES AFTER LONG-TERM WARMING EXPOSURE

William G. Rodríguez-Reillo

CHANGES IN SOIL MICROBIAL COMMUNITIES AFTER LONG-TERM WARMING EXPOSURE

A Dissertation Presented

by

WILLIAM GABRIEL RODRÍGUEZ-REILLO

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

SEPTEMBER 2019

Organismic and Evolutionary Biology

CHANGES IN SOIL MICROBIAL COMMUNITIES AFTER LONG-TERM WARMING EXPOSURE

A Dissertation Presented

by

WILLIAM GABRIEL RODRÍGUEZ-REILLO

Approved as to style and content by:

_____
Jeffrey L. Blanchard, Chair

_____
Courtney Babbitt, Member

_____
David Sela, Member

_____
Kristina Stinson, Member

_____
Paige Warren, Graduate Program Director
Organismic and Evolutionary Biology

DEDICATION

To my parents, William Rodriguez Arce and Carmen L. Reillo Batista. A quienes aún en la distancia me mantuvieron en sus oraciones. Los amo.

ACKNOWLEDGMENTS

I would like to thank Dr. Jeffrey L. Blanchard for opening the door of his lab, the guidance and support through my formation as a scientist in this academic journey. Also, I thank Dr. David Sela, Dr. Kristina Stinson, and Dr. Courtney Babbitt for serving on my dissertation committee and providing guidance in the most necessary moments.

None of this would be possible without help from the following founding sources, the Northeast Alliance for Graduate Education and the Professoriate (NEAGEP), Initiative for Maximizing Student Development (IMSD), National Science Foundation, Department of Energy and their Science Graduate Student Research Program. I want to thank the administrative personnel in the Organismic and Evolutionary Biology Program and IDGP because many complicate processes occur behind the scenes in those offices to maintain a functional program.

I also want to thank our collaborators Dr. Kristen DeAngelis, Dr. Susan Leschine, Dr. Vanessa Bailey, Dr. Jerry Melillo, and Dr. Serita Frey. My former lab mates and friends Kelly Haas and Lauren Alteio, thank you to make the lab a happy environment. Thanks to the OEB community, in special The Cohort, a small group with great heart and scientific knowledge. Special thanks to my scientific mentors and friends, Carolina Morell, Yadilette Rivera-Colón, Jesús Alvelo, and Wilbeth Morales. Last but not least, I am grateful of my old friends and support group Abner Rodríguez, Josué Vera, Raisa Canals, and Héctor Claudio. Thank you because the distance never constrained our interactions.

ABSTRACT

CHANGES IN SOIL MICROBIAL COMMUNITIES AFTER LONG-TERM WARMING EXPOSURE

SEPTEMBER 2019

WILLIAM G. RODRIGUEZ REILLO, B.A., UNIVERSITY OF PUERTO RICO ARECIBO

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Jeffrey L. Blanchard

Microbial metabolism is a key controller of ecosystem processes (e.g., carbon cycling). However, we are only starting to identify the molecular mechanisms and feedback in response to long-term warming. My dissertation integrates multi-omics techniques to capture changes in soil microbial communities after long-term warming exposure. The research projects leverage three warming sites (i.e., SWaN, Barre Woods, and Prospect Hill) located in Western Massachusetts at Harvard Forest. These sites provided a unique experimental setup to better understand microbes in response to long-term temperature change. For the three research projects, we delved into the (i) microbial biodiversity across all three warming sites, (ii) integration of soil carbon chemistry and metatranscriptomics at the Barre Woods site, (iii) and a time series of soil metatranscriptomes at the Prospect Hill site. Overall, these studies revealed a broader scope of changes occurring with long-term warming than anticipated. The warming treatment induced shifts in fungi groups and recalcitrant carbon decomposer bacteria. Changes in microbial functions involved metabolic pathways associated to biogeochemical and cellular stability as result of nutrient limitation. Further, our results provided new insights in microbial response to chronic temperature stress, suggested an ongoing change in community structure and function, and linked soil carbon decrease to cellular processes using high throughput

molecular techniques. This information will help to better understand interactions between microbial communities and the Earth's climate.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

TWO DECADES OF CHRONIC WARMING INTENSIFIES CHANGES IN BIODIVERSITY AND METABOLIC ACTIVITY AT HARVARD FOREST

1.1 Abstract

Soil microbial diversity is estimated to be immense, yet remains poorly understood due to challenges in the study of these complex communities. Nonetheless, microbial functions have been progressively recognized as key in driving biogeochemical cycles. Combining field experiments and high-throughput sequencing technologies represents a powerful tool to gain new insights of molecular functions controlling ecosystem-level processes. Here, we investigate soil community structure and function at three field warming experiments (SWaN, Barre Woods, and Prospect Hill) at the Harvard Forest Long-Term Ecological Research (LTER) site using metatranscriptomes to provide new insights into microbial response to simulated climate change. The experimental sites share a common treatment of being warmed continuously 5°C above ambient temperature. Adjacent plots, SWaN and Prospect Hill, were established 15 years apart which comparison of soils exposed to different durations of warming treatment. Bulk transcriptomes were extracted for soils in the organic horizon. Changes in cellular processes were measured by comparison of heated (+5°C) and control replicates. Results show that microbes consist of 26% eukaryotes and 73% bacteria, with the other 1% including viruses and archaea. Differences with the warming treatment included an increase in the bacteria to eukaryotes ratio, significant increases in Verrucomicrobia and Acidobacteria members, and decrease in fungi groups. Among other metabolic pathways, warming induced significant changes in proteins essential for structural integrity of the ribosome. Further, taxonomic

affiliation of ribosomal proteins has been linked to eukaryotic groups, including Basidiomycota and Ascomycota.

1.2 Introduction

Global surface temperature increased +0.70°C in 2017 relative to records from 1981 to 2010, marking a new record high (Arndt et al., 2018). Rising carbon dioxide (CO2) levels are mostly due to the burning of fossil fuels and the low rate of plant and bacterial belowground CO2 fixation. Terrestrial ecosystems serve as the largest global carbon reservoir, harboring 3000 Pg of carbon (Pries et al., 2017). However, it is uncertain how climate change can influence carbon turnover in soil (Schmidt et al., 2011). Therefore, it is necessary to better understand biological mechanisms in control of biogeochemical cycles, which in turn shape major ecosystem-level processes (Wieder et al., 2013).

Microorganisms are the most abundant and diverse form of life with up to 1012 species estimated on Earth (Locey and Lennon 2016). For instance, the prokaryotic diversity itself has been reported to account for up to 1.6 million species (Louca et al. 2019). Large efforts have been carried to link microbial diversity to ecosystem-level processes. The ability to capture bulk DNA and RNA from soil microbial communities circumvents previous culture-dependent constraints, including classic constraints of skewed isolation of readily culturable bacteria (Staley and Konopka, 1985). As high-throughput technologies become more accessible, our understanding of soil biodiversity will continue to expand. Investigations of soil microbial ecology using community approaches represent a powerful tool for biological discovery and formulate new hypotheses about the role of microorganisms in complex ecosystems.

Soil organic matter is the most readily available substrate to microbes. Changes in chemical properties allow carbon stocks to range from labile to stable compounds. Labile carbon

pools are subject to high rate of microbial decomposition with direct ecosystem-level impacts. In contrast, stable pools reflected a slow carbon turnover with low mineralization rate. For example, the woody tissue of trees is made up of cellulose, which can be mineralized by specialized microbes (Wei et al., 2009). Microbes play a key role in soil organic matter decay in distant ecosystems (Allison et al., 2010; Fierer and Jackson, 2006; Wild et al., 2014).

In Western Massachusetts, the Harvard Forest Long Term Ecological Research (LTER) site is home to three long-term warming experiments. These field sites share in common a 5°C increase above the ambient temperature using below ground heating cables. This warming effect has resulted in 33% carbon (C) losses and increased availability of nitrogen (N) to vegetation, degradation of the soil organic matter, shift in bacterial community composition, and decreased fungal biomass at the experimental sites (DeAngelis et al., 2015; Frey et al., 2008; Melillo, 2011; Melillo et al., 2017; Pisani et al., 2015). However, this is the first study to look at the soil metatranscriptomes across all three warming sites. Our objective was to determine changes in soil microbial biodiversity and functions in response to the warming treatment. We hypothesized microbial abundance shifts to access labile carbon over short-term warming exposure and/or changes in gene expression related to recalcitrant carbon substrates over long-term warming. Combining field experiments and high-throughput sequencing technologies represents a powerful tool to gain new insights of molecular functions controlling ecosystem-level processes.

1.3 Methods

1.3.1 Experimental Sites

Our soil warming experiments are located at the Harvard Forest LTER site in a mixed hardwood forest. The Soil Warming and Nitrogen addition (SW), Barre Woods (BW), and

Prospect Hill (PH) sites were started on 2006, 2003, and 1991, respectively. Dominant tree species includes maple (*Acer rubrum*), oak (*Quercus velutina*), and American beech (*Fagus grandifolia*). Atmospheric temperature ranges on average from –6°C in winter to 20°C in summer. Experimental treatment consists of a 5°C elevation in soil temperature relative to ambient control plots, using buried heating cables at 10cm depth (Peterjohn et al., 1993). The experimental treatment is shared across all three experimental sites.

1.3.2 Sample Collection and Sequencing

Soil samples were collected from all three warming experiments on October 2011. At the time of soil collection, the experimental sites had experienced the warming treatment effect for 5 years at SWaN, 8 years at Barre Woods (BW), and 20 years at Prospect Hill (PH). Soil samples (n=48) were collected from 3 sites, 2 soil horizons, 2 temperature treatments, and 4 replicates. On site, the organic samples were split into sterile bags and flash frozen in a ethanol and dry-ice bath. Samples were stored on dry ice for transportation to UMass and stored long-term at -80°C. Bulk RNA was extracted from 1 gram of soil using the MoBio PowerSoil kit for all mineral samples and an optimized CTAB protocol (DeAngelis et al., 2015). Samples were sent to the Joint Genome Institute (JGI) for library preparation and sequencing. Bacterial rRNA was depleted with the Illumina Ribo-Zero kit for all sequencing libraries. From the extracted RNA, cDNA libraries were prepared and sequenced on the Illumina HiSeq-2000 platform at the DOE JGI.

1.3.3 Bioinformatics

Sequenced libraries resulted in short (2x150 bp) pair-ended (PE) fragments. FastQC was employed to determine reads quality (Andrews, 2010). BBMap tools were used to remove adapters, trim, and filter low-quality reads (Bushnell, 2016). PE fragments were concatenated

using FLASH and rRNA sequences filtered with SortmeRNA (Kopylova et al., 2012; Magoč and Salzberg, 2011). Then, non-rRNA sequences were searched against the non-redundant NCBI database with DIAMOND and loaded into MEGAN for functional and taxonomic annotation (Buchfink et al., 2015; Huson et al., 2016).

1.3.4 Statistics

Statistical analyses were done using R and implemented in RStudio (R Core Team and R Development Core Team, 2008; RStudio Team, 2016). The R package edgeR was employed to determine enrichment of features (Lun et al., 2016; Zhou et al., 2013). Comparisons included by site (3 Sites x 2 Treatments x 4 replicates=24) and all three sites combined. Pooling all three sites together improves statistical power by increasing the total number of samples. Trimmed Mean of M-values (TMM) method was used to normalize genomic counts. Normalization factors were calculated for each experimental site and all sites together to reduce variance. Genomic features with significant treatment effect considered comparison with an FDR≤0.1.R package ggplot2 and Gephi software were used to visualize genomic data sets (Bastian and Heymann, 2009; Wickham, 2009).

1.3.5 Data Availability

Raw sequences are publicly available at https://img.jgi.doe.gov/m. IMG accession numbers are included in the Table 1.3.

1.4 Results and Discussion

Soil metatranscriptomes were compared to determine the (i) treatment effect within sites (SW (5yrs); BW (8yrs); and PH (20yrs)) and all sites together and (ii) long-term warming effect. The latter is based on comparison between the SW-5yrs and PH-20yrs heated (+5°C) soils

relative to the control. This study is centric to soil metatranscriptomes, but briefly compared metagenomes at the domain taxonomic resolution. Soil metagenomes were further described in a previous publication (Pold et al., 2016). Sequencing of metatranscriptomes extracted from soils collected at the Harvard Forest LTER produced a total of 21,305,735 protein coding reads (or mRNA reads) obtained from 24 organic soil samples (Table 1.1). Sequence reads were assigned to 1,667 taxonomic groups and 5,534 functional features. At the domain resolution, taxonomic breakdown resulted in 73% bacteria, 26% eukaryotes, and 1% others (including Archaea 0.07% and Viruses 0.37%) (Figure 1.1).

Soil eukaryotes were better represented by the metatranscriptomes than the metagenomes, representing 26% of the total community in RNA as compared to 1.5% in the DNA (Figure 1.1). From a biological standpoint, it is expected for soil eukaryotes to be present and play an active role in the forest ecosystem (Wiesmeier et al., 2019). Although our study targeted all messenger RNA (mRNA) sequences, other metatranscriptomes studies using polyadenylated mRNA captured the rich functional diversity and contribution within soil eukaryotes (Bailly et al., 2007; Lehembre et al., 2013). Much more studies have been done to target fungi, which represent a high economic impact (Kennedy and Stajich, 2015; Ohm et al., 2012).

At the Harvard Forest sites, the most abundant eukaryotic groups include fungi (e.g., Ascomycota, Basidiomycota), green plants (e.g., Streptophyta), and invertebrates (e.g., Arthropods), which are essential components of the soil food web (Figure 1.1). Transcripts related to bacteria accounted for 73% of all protein sequences. Most abundant groups included Proteobacteria, Acidobacteria, Actinobacteria, Planctomycetes, Verrucomicrobia, and Bacteroidetes (Figure 1.1). These groups have been described as ubiquitous soil-dwelling

bacteria based on genetic markers (Janssen, 2006). More recently new efforts categorized some of this bacterial groups (e.g., Acidobacteria, Actinobacteria and beta-proteobacteria) based on their lifestyle strategies, i.e. copiotroph or oligotroph (Fierer et al., 2007). For instance, previous studies in the oldest Harvard Forest site reported that long-term warming favors oligotrophic conditions and microbial communities (DeAngelis et al., 2015; Melillo et al., 2017). Under this lifestyle classification context, our warming treatment resulted in low nutrient available to soil bacteria as labile carbon become limited and enhanced over time.

The experimental warming treatment resulted in an overall increase in the ratio of bacteria to fungi, which is more pronounced at the two oldest sites, Barre Woods (8 yrs of warming) and Prospect Hill (20 yrs of warming) (Figure 1.2). A similar community response has been reported previously for which quantitative PCR show a decrease in fungi and the 16S ribosomal RNA marker gene increase for some bacterial groups (DeAngelis et al., 2015). Other studies also refer to an imbalance of the bacteria to fungi biomass by predation of protist on fungi (Ruess et al., 1999; Schroter et al., 2003; Sjursen et al., 2005). Further, the interaction of fungi and their predators can alter feedbacks to ecosystem-level functions (Crowther et al., 2015). In our field sites, the temperature increase may serve as mechanism to recruit more fungi predators, which can explain the observed decrease in fungal biomass. This potential recruitment of predators is conceivable as atmospheric temperature falls during the time of soil collection. Warming response across taxonomic ranks were presented as well (Table 1.2).

The microbial response to warming was variable across sites, however, soil eukaryotes across taxonomic levels consistently decreased in abundance with warming (Figure 1.3). This fungi response to warming has been reported before using phospholipid fatty acid (PLFA) profiles (Frey et al., 2008). Also, soil metagenomes reported a decrease of carbohydrates

degrading enzymes from eukaryotes and a dual shift in carbohydrate degradative enzymes (Pold et al., 2016). In turn, this study captured the activity of soil eukaryotes and other soil microbes using protein coding sequences. Most eukaryotic sequences were linked to fungi, with saprotrophs and pathogens representing 15% of transcripts and showing the most sensitivity to warming. Warming effect within fungi resulted in significant changes of Dothideomycetes, Leotiomycetes,Sordariomycetes, Mortierellomycotina classes (Figure 1.3). This soil-dwelling fungi classes showed a significant treatment effect in the oldest site, Prospect Hill (20-yrs). At lower taxonomic rank, the Gloniaceae fungal family was the most represented group within Dethideomycetes. Members of this family are known to include ectomycorrhizae and saprophytic fungi (Spatafora et al., 2012). The Leotiomycetes class significantly decreased with warming and most transcripts linked to Myxotrichaceae and Pseudeurotiaceae families. Leotiomycetes group is notorious plant endophytes, which allow members to be present in large range of environments (Wang et al., 2006). Moreover, this group of endophytic fungi has been implicated to control carbon storage (Tong et al., 2017). The Sordariomycetes class decrease with most transcripts linked to Hypocreaceae family. Members within this fungal class has been described as pathogens, endophytes, and soil organic matter degraders (Uroz et al., 2013). Mortierellomycotina group are commonly associated to the plant rhizosphere and host for endosymbiotic bacteria (e.g., Betaproteobacteria and Mollicutes) (Bonfante and Desirò, 2017; Desirò et al., 2018; Summerbell, 2005). This is of interest as potential endosymbiotic bacteria has been detected in the metatranscriptomes and changed in abundance with warming (see below).

Bacteria and Eukaryotes responded differently to the warming effect. As opposed to fungi, most bacteria groups increased with the warming treatment. The dominant bacterial groups reflected a significant treatment effect with Verrucomicrobiaceae, Gemmatimonadetes,

and Blastocatellia classes increasing in abundance with warming (Figure 1.3). These groups have been described as oligotrophic bacteria, which typically grow slowly and inhabit nutrient poor environmental niches (Bergmann et al., 2011; Cederlund et al., 2014; Eichorst et al., 2011; Fierer et al., 2012; Pascual et al., 2015; Smit et al., 2001). When the SW-5y and PH-20y heated plots were compared, the long-term warming effect reflected an increase in abundance of soil bacteria at the field site longer exposed to warming (PH-20y) (Figure 1.4). From a trophic lifestyle perspective, we expected to observe an oligotrophic community establish in the PH-20y site relative to SW-5y as carbon substrates became limited after long-term warming exposure. In both sites, we observed bacterial groups associated to oligotrophic conditions (e.g., Actinobacteria, Gammaproteobacteria, etc.). However, the overall abundance of bacterial groups was greater at the PH-20y site.

Warming induced shifts in microbial processes related to biogeochemical cycles. Significant shifts in transcriptional abundance of ecosystem-level processes included sulfur oxidation, nitrate and nitrite ammonification, hemin transport system, iron acquisition, and benzoate degradation (Figure 1.4). Metabolism related to sulfur oxidation increase in abundance with warming at the BW-8y and PH-20 sites. Sulfur oxidation is a key metabolic process employed by chemolithotrophic bacteria for energetic budgets during respiration (Berben et al., 2019). In terrestrial ecosystem, this sulfur-oxidizing bacteria has been linked to Alphaproteobacteria and betaproteobacteria and isolated from bulk soil and the rhizosphere (Graff and Stubner, 2003; Wohl et al., 2004). Further, the microbial transformation of sulfur compounds (e.g., sulfite and sulfate) interconnect the sulfur cycle with other biogeochemical cycles (e.g., nitrogen, iron, and manganese) (Berben et al., 2019). There are three main pathways employed by sulfur-oxidizing bacteria (Ghosh and Dam, 2009). This includes the SOX pathway, which is mediated by a thiosulfate-oxidizing multi-enzyme complex. At the Harvard

Forest, the warming treatment has resulted in significant increase of transcripts encoding for the sulfite dehydrogenase cytochrome subunit SoxD and the sulfur oxidation molybdopterin C proteins.

The experimental treatment also resulted in decrease of the nitrate and nitrite ammonification pathway at BW-8y site (Figure 1.4). This metabolic process is essential for microbial energy and vascular plants. After 4 years of warming increase in Harvard Forest, the treatment stimulated root exudation and promoted transformation of soil nitrogen compounds in the rhizosphere (Yin et al., 2013). There are three main mechanisms that bacteria employ to assimilate nitrate and nitrite, including ABC transporters driven by ATP hydrolysis, secondary transporters based on a proton gradient, and the NarK transport system (Moir and Wood, 2001). The NarK system is analogous to the secondary transport mechanism and involved in denitrification processes. At the BW-8y site, all five annotated transcripts within the nitrate and nitrite ammonification pathway decreased with warming, however, not significantly affected at the gene resolution. The most abundant transcript encoded for the nitrate/nitrite transporter family NarK (pval=0.007 or after multiple test correction FDR=0.2). At our own BW site, it has been reported that the warming treatment increase inorganic nitrogen availability in soil to vegetation (Melillo et al., 2011). In the same site, the observed decrease in transcripts related to the nitrate and nitrite pathway with warming may result as inorganic nitrogen is readily available in the heated plots and cellular investment (e.g., ammonification) occurred at the control plots to compensate for the limitation.

At the PH-20y, microbial processes related to iron metabolism depicted significant differences in the iron acquisition and hemin transport system pathways (Figure 1.4). Iron is an

essential element in aerobic bacteria for energy synthesis and heme production (Neilands, 1995). Heme play an important role in electron transfer, redox activity, and function of conserved enzymes (e.g., nitric oxide) (Choby and Skaar, 2016; Faller et al., 2007). The warming effect on the hemin transport system was associated to the outer membrane receptor protein of iron transport. Based on the functional annotation, this protein has been detected in several bacterial groups, including Bacteroidetes and Betaproteobacteria. Both bacterial groups contain endosymbionts of insects and fungi (Bertaux et al., 2005; Gruwell et al., 2007; Sharma et al., 2008). The Burkholderiales decreased with the warming treatment, with transcripts decreasing 2% (control 7.7%; heated 5.3%) on average across the experimental sites. Burkholderiales has been described previously as endofungal bacteria (Partida-Martinez and Hertweck, 2005). Interestingly, the Burkholderiales and Fungi groups decreased both with the experimental treatment. Therefore, the warming effect can influence pathogenetic response and iron metabolism.

Another significant effect with warming was detected on the increase of iron acquisition pathway linked to the TonB dependent receptor gene. This gene is among the most abundant mRNA and significantly increase with warming. The TonB system has been reported to be involved in metals (e.g., iron) uptake and aromatic metabolism (i.e., uptake and degradation) (Jordan et al., 2013; Miller et al., 2010). In this study, transcripts encoding for TonB dependent receptor were mostly linked to Acidobacteria, whose transcripts represent 6% for this gene. Acidobacteria as a phylum increased in response to warming as measured by 16S rDNA (DeAngelis et al., 2015). Overall, the Acidobacteria transcript abundance was not significantly different, however, specific genes like TonB associated to these phyla were more abundant. This highlights the significant role of this bacterial group to drive major ecosystem-level processes, including degradation of aromatic compounds.

A small set of metabolic pathways doubled in abundance in response to warming (Figure 1.5). Most of these pathways were linked to carbohydrates and protein metabolism with a significant treatment effect recorded in the oldest warming sites (BW-8y and PH-20y). Cellular processes were linked to melibiose utilization, regulatory intramembrane proteolysis, mycofactocin system, ribosome small subunit (SSU) mitochondrial, and ribosome activity modulation. The melibiose utilization and ribosome activity modulation pathways were mostly associated to bacteria, whereas the regulatory intermembrane proteolysis and ribosome SSU mitochondrial pathways linked exclusively to eukaryotes (Figure 1.6). The pathways specific to eukaryotes were linked to Basidiomycota and Ascomycota fungal groups.

Table 1.1: **Read abundance per soil metatranscriptomes. The columns show values for the sequenced sample (Sample ID), unprocessed reads (Raw reads), filtered and trimmed (Cleaned reads), concatenated pair-ended reads (Merged reads), ribosomal RNA reads as identified with sortmeRNA (rRNA reads), reads without hits with sortmeRNA (non-rRNA reads), and protein coding reads (mRNA reads).**

Sample ID	Raw reads	Cleaned reads	Merged reads	rRNA reads	non-rRNA reads	mRNA reads
SWNH8O	28,394,128	27,767,995	22,609,438	1,114,255	21,495,183	4,698,850
SWNH22O	22,899,814	22,139,832	13,554,354	11,394,205	2,160,149	511,594
SWNH18O	19,950,891	19,332,394	14,554,314	12,064,607	2,489,707	538,909
SWNH13O	22,019,741	21,267,061	15,581,401	11,491,015	4,090,386	760,988
SWNC24O	21,232,603	20,459,337	16,267,573	13,767,299	2,500,274	505,058
SWNC20O	21,184,173	20,364,085	15,620,163	11,568,937	4,051,226	665,331
SWNC19O	22,969,898	22,133,201	14,668,081	12,666,311	2,001,770	496,539
SWNC1O	21,217,402	18,664,412	12,924,204	10,559,433	2,364,771	452,910
PHH8O	21,345,791	20,408,514	15,285,510	11,143,642	4,141,868	710,496
PHH6O	23,318,136	22,048,297	16,424,286	11,371,089	5,053,197	851,621
PHH15O	22,391,897	21,327,642	16,589,841	11,729,858	4,859,983	813,582
PHH12O	60,948,309	57,685,049	38,403,481	25,556,132	12,847,349	1,943,205
PHDC5O	19,718,387	18,781,861	14,514,427	9,897,780	4,616,647	726,656
PHDC3O	22,742,166	21,723,875	16,647,997	13,606,915	3,041,082	644,179
PHDC13O	29,757,090	28,095,773	21,665,222	17,294,562	4,370,660	1,036,906
PHDC10O	20,932,386	20,106,232	15,244,691	12,127,071	3,117,620	670,425
BWH4O	22,880,148	21,841,144	16,816,005	12,006,986	4,809,019	991,347
BWH30O	27,235,071	25,575,314	19,697,961	14,377,499	5,320,462	888,935
BWH19O	29,014,894	27,600,128	21,333,707	18,450,500	2,883,207	529,581
BWH17O	27,398,086	25,984,681	20,653,211	17,254,737	3,398,474	723,074
BWC4O	23,934,131	23,011,105	17,994,286	15,278,823	2,715,463	399,695
BWC30O	25,320,446	24,396,589	18,749,646	16,589,539	2,160,107	495,620
BWC19O	28,743,260	27,191,286	20,677,433	17,297,784	3,379,649	706,637
BWC17O	26,242,676	24,971,396	19,560,560	16,109,538	3,451,022	543,594
Total	611,791,524	582,877,203	436,037,792	324,718,517	111,319,275	21,305,732

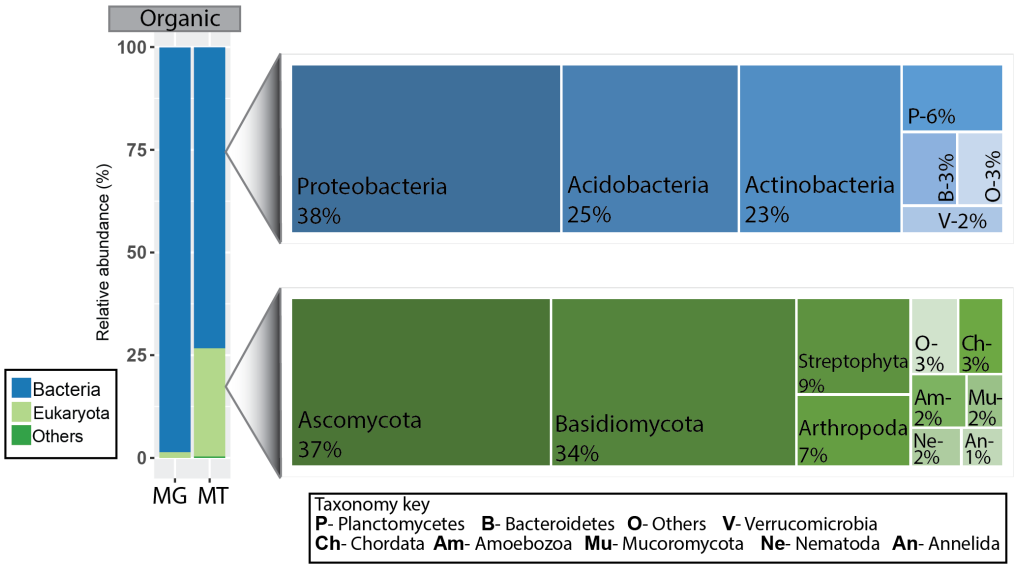

Figure 1.1: Soil biodiversity at the Harvard Forest in the organic horizon. Stacked bar plots show relative abundance of microorganisms at domain resolution for metagenomics (MG) and metatranscriptomics (MT) datasets. Treemaps in the right side expand into the bacteria and eukaryotes groups.

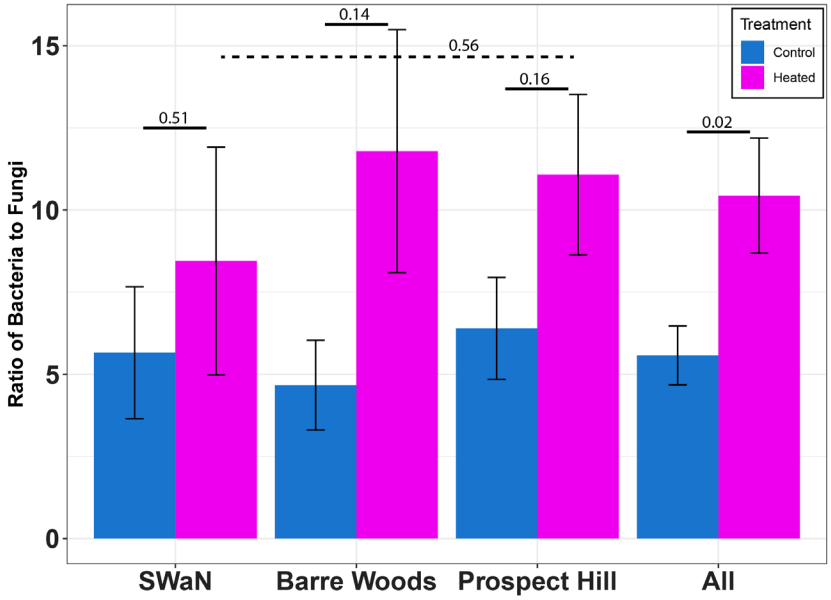

Figure 1.2: Warming effect in the abundance of bacteria relative to fungi mRNAs in the organic horizon for the three warming sites (SWaN, Barre Woods, and Prospect Hill) and all sites combined (All). The bar plot shows the experimental sites on the x-axis, ratio of Bacteria to Fungi on the y-axis, standard error as whiskers, and above the p-value (t-test) for the treatment effect. The dotted line compares heated plots for the adjacent warming sites SWaN and Prospect Hill.

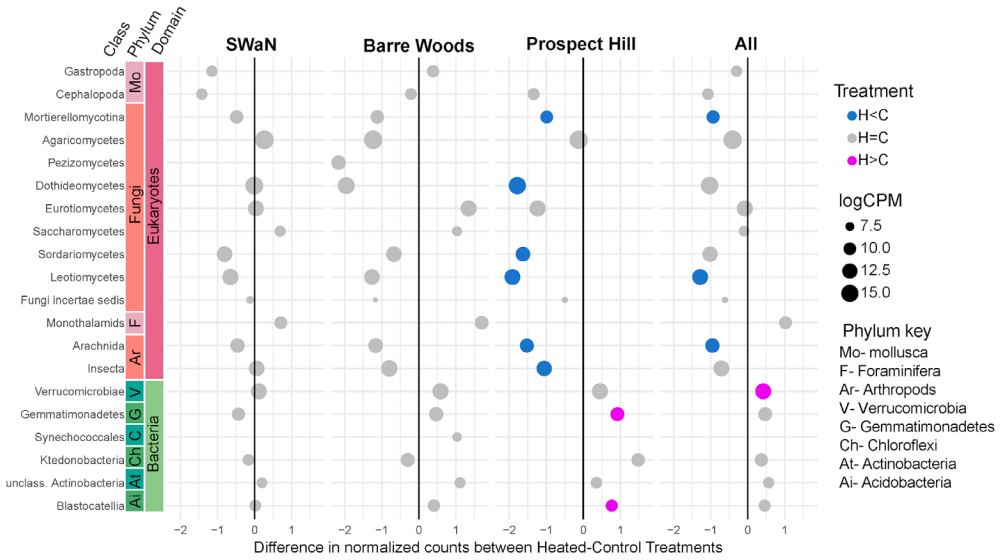

Figure 1.3: Changes in soil biodiversity with the warming treatment. Bubble plot depicts selected taxonomic classes groups with fold change above relative changes with the warming treatment effect. Taxonomic groups with domain, phylum, and class classifications on x-axis.

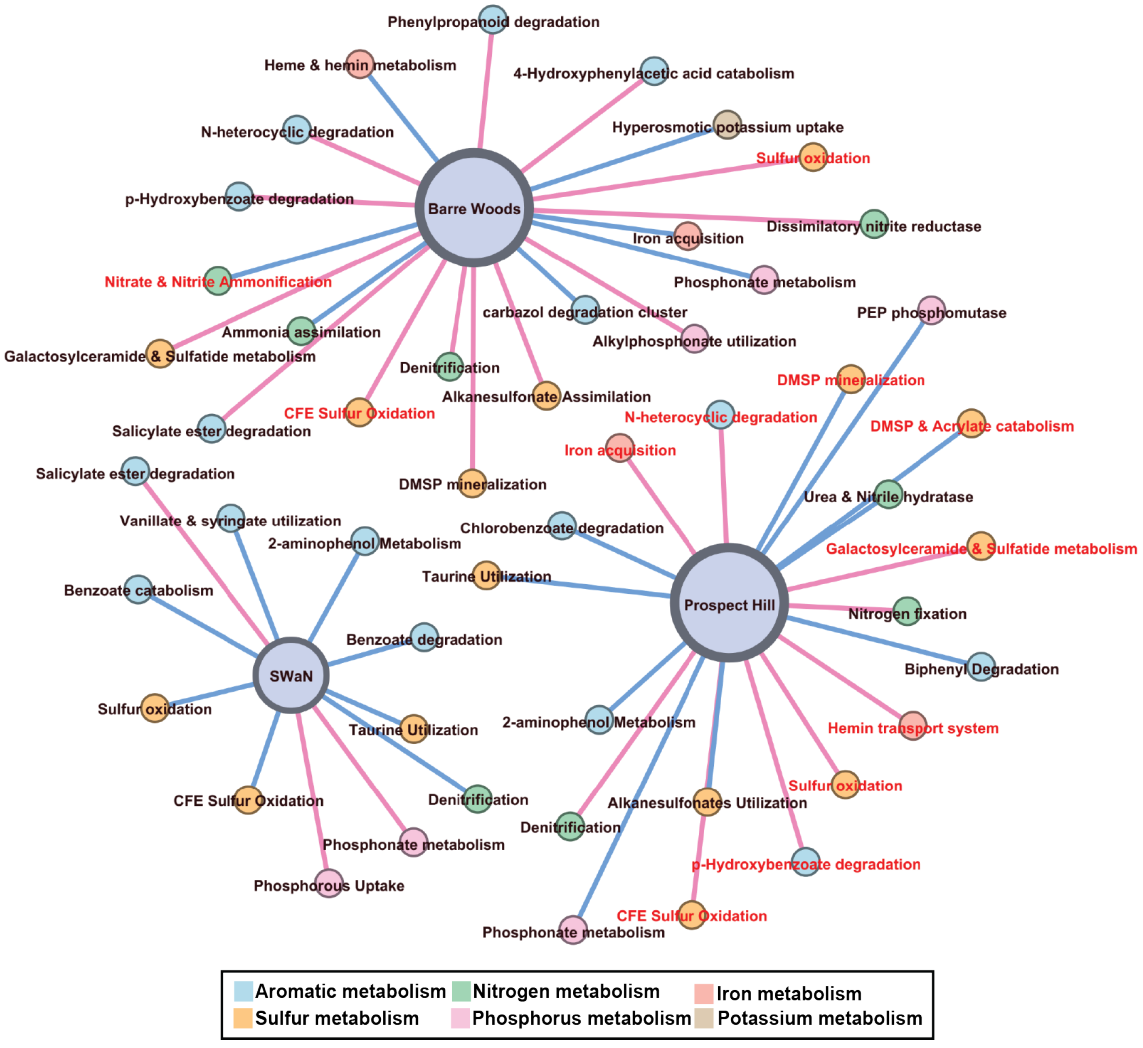

Figure 1.4: Microbial metabolism associated to biogeochemical processes at the Harvard Forest warming sites. Large nodes in purple depict the three experimental sites (SWaN-5yrs, BW-8yrs, and PH-20yrs) with connecting nodes as metabolic pathways. Pathway-nodes were colored based on metabolism according to the key above, and red font depict pathways with significant (FDR≤0.1) treatment effect. Edges capture the treatment effect of transcriptional abundance; i.e. red means increase and blue decrease in abundance on metabolic pathways. pathways filtered to include processes for which the treatment effect (log2 fold change) is situated outside the interquartile range.

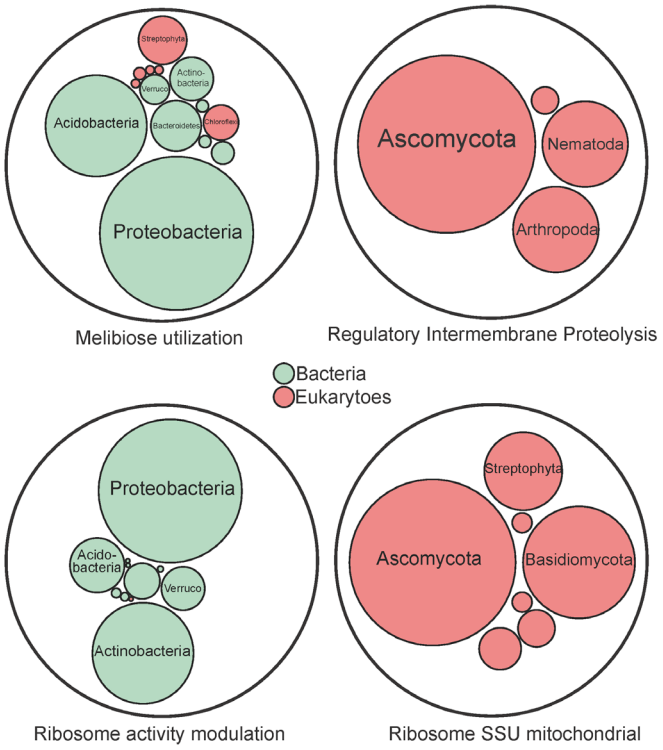

Figure 1.5: Taxonomic affiliation of metabolic pathways with significant treatment effect in the organic horizon. Large circles represent the metabolic pathways which contain smaller circles that represent the taxonomic groups.

Table 1.2: Overall shift in taxonomic groups across taxonomic resolution. Warming effect is capture with colors, red means increase, and blue means decrease in abundance of the taxon. Statistical significance recorded within parenthesis, non-significant (n.s.) or comparison id for significant taxon (FDR≤0.1). Comparisons include each site (SWaN (SW), Barre Woods (BW), Prospect Hill (PH)) and all together (All).

Organic				
Domain	**Phylum**	**Class**	**Order**	**Family**
Bacteria (n.s.)	Acidobacteria (n.s.)	Blastocatellia (PH)	Environmental samples Acidobacteria (PH)	-
	Actinobacteria <phylum> (n.s.)	Actinobacteria <class> (n.s.)	Corynebacteriales (PH)	Mycobacteriaceae (PH)
			Micrococcales (PH)	Microbacteriaceae (BW; All)
	Gemmatimonadetes <phylum> (PH)	Gemmatimonadetes <class> (PH)	Gammatimonadetes (PH)	-
	Proteobacteria (n.s.)	Gammaproteobacteria (n.s.)	Enterobacterales (PH)	Enterobacteriaceae (BW; All)
	Verrucomicrobia (n.s.)	Verrucomicrobiae (PH)	Verrucomicrobiales (All)	Verrucomicrobia subdivision 3 (All)
Eukaryotes (BW)	Arthropoda (PH; All)	Arachnida (PH; All)	Araneae (PH; All)	-
			Astigmata (PH)**	-
		Insecta (PH)	Hymenoptera (PH)	-
	Ascomycota (PH; All)	Dethideomycetes (PH)	Pleosporomycetidae incertae sedis (PH)	-
		Eurotiomycetes (n.s.)	Eurotiales (PH)	-
		Leoteomycetes (PH; All)	Helotiales (PH)	-
			Leotiomycetes incertae sedis (PH)	Myxotrichaceae (PH)
		Sordariomycetes (PH)	Hypocreales (PH)	-
	Basidiomycota (BW)	Agaricomycetes (n.s.)	Agaricales (n.s.)	Tricholomataceae (BW)
	Chlorophyta (BW; All)	-	-	-
	Chordata (PH; All)	-	-	-
	Foraminifera (BW)	-	-	-
	Mucoromycota (n.s.)	Mortierellomycotina (PH; All)*	Mortierellales (PH)	-
Viruses (BW)	-	-	-	-

Table 1.3: IMG accession numbers for metatranscriptomes.

SampleID	IMG ID
BWC17O	3300002650
BWC19O	3300002659
BWC30O	3300002648
BWC4O	3300002649
PHDC10O	3300002664
PHDC13O	3300002672
PHDC3O	3300002653
PHDC5O	3300002652
SWNC1O	3300002645
SWNC19O	3300002662
SWNC20O	3300002660
SWNC24O	3300002651
BWH17O	3300002658
BWH19O	3300002655
BWH30O	3300002673
BWH4O	3300002668
PHH12O	3300003551
PHH15O	3300002656
PHH6O	3300002667
PHH8O	3300002661
SWNH13O	3300002663
SWNH18O	3300002654
SWNH22O	3300002657
SWNH8O	3300004121

CHAPTER 2

INTEGRATION OF SOIL METATRANSCRIPTOMES AND MASS SPECTROMETRY AFTER 15-YEAR OF CHRONIC WARMING

2.1 Abstract

In terrestrial ecosystems, soils contain about three times more carbon than is in the atmosphere. Microbial metabolism is a key controller of soil carbon dynamics and other soil biogeochemical transformations. However, we are only starting to identify the molecular processes and changes involved to the soil carbon pools in response to long-term warming. In our long-term soil warming experiment in the Barre Woods tract at Harvard Forest, a 5o elevation in soil temperature has increased respiration rates and the loss of soil carbon to the atmosphere. In order to link microbial metabolism with soil biogeochemistry, we measured cellular metabolism and soil organic matter chemistry using both next generation sequencing and high-resolution mass spectrometry (MS). The high resolving power and measurement accuracy of masses detected with MS enabled us to identify 16,814 carbon molecules, including 58.7% with assigned molecular formula and 15.2% with hits to the KEGG database. Changes associated to lignin, hydrocarbons, and lipids reveal significant differences with the warming treatment. Microbial diversity also shows shifts in response to the temperature treatment, including metabolic changes associated to degradation of benzoic compounds, denitrification and ammonification, sulfur assimilation, and potassium uptake. Other metabolic changes associated with cellular processes such as dormancy and sporulation also increased with the experimental treatment. Taxonomic distribution differed between the heated soils and control, revealing a shift in the abundance of fungi, archaea viruses, and bacterial phyla in response to long-term warming. Network analysis integrated both MS and metatranscriptomes providing new insights on microbial carbon transformations. This integrative approach provides a

framework to inform soil respiration models that can be incorporated into ecosystem and biogeochemical processes.

2.2 Introduction

Large efforts have contributed to the understanding of carbon feedbacks in response to change in global temperature. The advancement has been constrained by limitations in our basic understanding of ecosystem processes. For instance, soil carbon reservoirs serve as a carbon sink, sequestering carbon from the atmosphere, but can switch to a source in response to climate change, thereby releasing carbon with the potential to contribute to further climatic shift. However, the magnitude and rate to which this can occur is not clear (Heimann and Reichstein, 2008). High complexity in soil reservoirs resides as mixture of compounds with variable turnover (Trumbore, 1997). These factors shape carbon pools and influence nutrients available to microorganisms (Dungait et al., 2012). Thereby influencing microbial communities and function (Carney et al., 2007; DeAngelis et al., 2015; Eichorst et al., 2011).

Soil organic matter encompasses a broad spectrum of carbon substrates ranging from labile to recalcitrant compounds found in soil. Further, the heterogeneity among chemical properties within the C pools in the forest soil are not well-defined (Kleber et al., 2011). New high-resolution mass spectrometry methods (i.e., FTICR-MS) however reveal the molecular structure of constituents of the SOM (Simon et al., 2018; Tfaily et al., 2015). Some studies have shown the interconnection between soil chemistry and microbial diversity when coupled with amplicon and DNA probes (Ward et al., 2017; Wu et al., 2018).

In our 15-year experiment, warming has resulted in increased respiration rates and loss of soil carbon to the atmosphere (Melillo et al., 2011). Here, we integrated soil chemistry and

microbial biodiversity to gain insights of carbon transformation linked to microbes. Soil metatranscriptomes captured microbial metabolism while FTICR-MS detected carbon molecules in the forest soil. Integration of data across these approaches depicted changes in chemical properties, microbial biodiversity, and carbon-enzyme groups acting on SOM transformation.

2.3 Methods

2.3.1 Sample collection and RNA extraction

Soils were collected on May 24, 2017 from the Barre Woods long-term experimental warming plots located at the Harvard Forest Long Term Ecological Research (LTER) site in Petersham, MA. Fourteen soil cores were collected from subplots within the larger 30x30 meter plots. The experimental site has been further explained in previous publication (Melillo, 2002). Soil collection and RNA extractions were also explained in a former publication (Schulz et al., 2018). Briefly, soils were separated into organic and mineral horizons by visual inspection and total RNA extracted from 28 samples (mineral; n=14 and organic; n=14) using the RNeasy PowerSoil RNA extraction kit (QIAGEN). cDNA libraries were prepared and sequenced on the Illumina NextSeq platform at the Joint Genome Institute (JGI).

2.3.2 FT-ICR-MS solvent extraction and data acquisition

Mass spectrometry analysis was carried on the same soil samples (n=28) collected for RNA extractions. For the solvent extraction procedure 500 mg of soil was weighed out into 2mL glass vials. We then performed a modified Folch extraction, protocol 6, by first extracting the soils with 1mL of MilliQ water (Tfaily et al., 2017a). This chemical extraction protocol combined with Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) is used to probe differences in metabolites among our samples. A 12 Tesla Bruker SolariX (Bruker solariX,

Billerica, MA) FTICR-MS was outfitted with a standard electrospray ionization (ESI) interface and operated in negative mode. Samples were directly infused at a flow rate of 3.0 µL min-1 into the mass spectrometer using a home built automated Pal Autosampler (HTX technologies) coupled with Agilent 1200 series pumps (Agilent Technologies) followed by two off line blanks. Experimental conditions were as follows: needle voltage, +4.2 kV; Q1 set to 50 m/z; and the heated resistively coated glass capillary operated at 180 °C. Data were collected by co-adding 144 scans from 100 m/z to 900 m/z at 4M and the ion accumulation time was optimized for each sample.

2.3.3 FT-ICR-MS data processing

One hundred forty-four individual scans were averaged for each sample and internally calibrated using an organic matter homologous series separated by 14 Da (–CH2 groups). The mass measurement accuracy was less than 1 ppm for singly charged ions across a broad m/z range (100 - 900 m/z). Data Analysis software (BrukerDaltonik, v.4.2) was used to convert raw spectra to a list of m/z values applying FTMS peak picker module with a signal-to-noise ratio (S/N) threshold set to 7 and absolute intensity threshold to the default value of 100. Chemical formulae were then assigned using in-house software following the Compound Identification Algorithm (CIA), proposed by Kujawinski and Behn, modified by Minor et al., and described in Tolic et al. (Kujawinski and Behn, 2006; Minor et al., 2012; Tfaily et al., 2017b). Chemical formulae were assigned based on the following criteria: S/N >7, and mass measurement error < 0.5 ppm, taking into consideration the presence of C, H, O, N, S and P and excluding other elements.

2.3.4 Bioinformatics and Statistics

Raw pair-ended (2x150) sequences were merged and cleaned (i.e., filter and trimmed) using Flash and BBDuk, respectively (Bushnell, 2016; Magoč and Salzberg, 2011). FastQC tool was used to assess the quality of sequences (Andrews, 2010). Then, SortmeRNA was employed to remove ribosomal RNA sequences from further analyses (Kopylova et al., 2012). Retained sequences were searched against the non-redundant NCBI database with Diamond and output loaded into MEGAN for functional and taxonomic annotations (Buchfink et al., 2015; Huson et al., 2016). Statistics on both FTICR-MS and metatranscriptomes were performed in R environment (R Core Team and R Development Core Team, 2008). Data visualizations were performed in R with packages ggplot2 and plotly and Gephi (Bastian and Heymann, 2009; Chen et al., 2015; Wickham, 2009).

2.3.5 Data Integration

Soil chemistry and microbial profiles were integrated using the Kyoto Encyclopedia of Genes and Genomes (KEGG) as metabolic framework (Okuda et al., 2008). The metabolic framework provides ecological information of microbial metabolism associated to enzymes and biochemical compounds. Further, this link the enzymes and compounds to ecosystem-level processes. For example, the nitrite reductase (KEGG EC:1.7.2.1) enzyme catalyzed the chemical compounds nitric oxide (KEGG ID: CPD:C00533) to nitrite (KEGG ID: C00088), which play an important role in the nitrogen cycle (KEGG pathway: Nitrogen metabolism). Based on the FTICR-MS dataset, the biochemical compounds detected with assigned molecular formula were searched for hits against the KEGG compounds. In turn, the metatranscriptomes dataset informed about the enzymes for which the enzyme commission (EC) identifier was used as

reference against the KEGG metabolic pathways. The manually curated SEED functional annotation was used as reference for enzymes. Since more than one enzyme can act on more than one compound and vice versa, our approach parsed chemical reactions to account for such interactions. Therefore, the visualized integrative network contains enzyme-nodes shared with multiple carbon compounds. When both FTICR-MS and metatranscriptomes datasets are taken together, this enabled our integrative approach to inform of chemical reactions.

2.3.6 Data Availability

Metatranscriptomics sequences are publicly available at https://img.jgi.doe.gov/m, IMG accession numbers included in Supplementary Table 2.2.

2.4 Results

2.4.1 Soil Metatranscriptomes

Sequencing of metatranscriptomes extracted from soils collected at the Barre Woods site produced a total of 184,499,712 protein coding reads obtained from across 28 soil samples with 7 biological replicates for each soil layer in each condition. Sequences were assigned to 1,483 taxonomic groups and 6,241 functional features. Abundance of some microbes was significantly different with the warming effect (Figure 2.1). Most taxonomic groups with significant treatment effect occurred at the class rank. However, the treatment resulted in significant changes across taxonomic ranks (Table 2.1). Most transcripts associated to prokaryotes (i.e., archaea and bacteria) increased in abundance with the experimental treatment. In contrast, the eukaryotes and viruses decreased in abundance with the temperature effect. As in the previous chapter, the abundance of fungi to bacteria decreased in response to long-term warming. Transcripts associated to viruses significantly decrease with the experimental treatment in both

soil layers (Figure 2.1). Most viral transcripts were linked to single stranded RNA (ssRNA) and positive-stranded sequences. From the diversity metrics tested, the Pielou's evenness metric depicted a significant warming increase in the organic horizon. Fungi dominated the organic layer relative to the mineral. Both soil horizons reflected a significant decrease in the Leotiomycetes and Agaricomycetes fungal groups and Sordariomycetes only in the mineral horizon (Figure 2.1). The temperature effect altered abundance of bacterial members within four major phyla, including Actinobacteria, Chloroflexi, Acidobacteria, and Proteobacteria (Table 2.1). Most bacterial groups with significant treatment effect increased in abundance in both soil horizons. However, members within Betaproteobacteria and Gammaproteobacteria groups showed a significant decrease with warming at the organic horizon. More specifically, the Betaproteobacteria group was associated to the Burkholderiaceae family and Gamma-proteobacteria linked to Moraxellaceae and Rhodanobacteraceae families. The increase in bacteria encompass the Actinobacteria and Chloroflexi phyla, which both major groups reflected a similar treatment effect in both soil horizons (Figure 2.1). A significant difference in abundance was recorded for Mycobacterium in the both soil horizons and Thermoleophilia in the mineral within the Actinobacteria phyla. In turn, the Ktedonobacteria group accounted for the observed changes in the Chloroflexi phyla.

Changes in functional processes were linked to several biogeochemical processes (Figure 2.2). This included changes in carbohydrates, aromatics, sulfur (S), nitrogen (N), phosphorus (P), and potassium (K) metabolism. For instance, the transcriptional abundance of carbon metabolism, respiration, and cellular processes increased in abundance with the temperature effect in the organic horizon (Figure 2.3). In contrast, S and P metabolism in conjunction with

viral related pathways decreased in abundance. Metabolic pathways annotated in the mineral horizon showed an overall increase with the treatment effect (Figure 2.4).

Metabolic pathways were linked to three major taxonomic groups in both soil horizons, including Acidobacteria, Actinobacteria, and Alpha-proteobacteria. Trehalose metabolism was mostly associated to Bradyrhizobiales and Beijerinckiaceae groups in both soil horizons. Nitrogen metabolism was linked to decrease in ammonia assimilation and increase in denitrification processes. At the gene resolution the respiratory nitrate reductase alpha and beta chains were linked to the Bradyrhizobiaceae group. The Thaumarchaeota archeal phyla was also associated to the nitrogen cycle. This group was linked to the copper-containing nitrite reductase gene. Further, members within Thaumarchaeota contain ammonia-oxidizers. Changes in aromatic degradation included toluene, quinate, and n-Phenylalkanoic acids (Figure 2.2). The taxonomic affiliation among these aromatic pathways were mostly associated to Actinobacteria and Alpha-proteobacteria phyla in both soil layers. Burkholderiaceae was associated to an increase in benzoate degradation activity with the warming treatment.

Changes in abundance of the carbohydrates-active enymes (CAZy) were detected in the metatranscriptome comparison (Figure 2.5). Abundance of most significant CAZy families decrease in abundance with the temperature effect. Transcripts involved to the Polysaccharide Lyase (PL) family 14 (subgroups 3,4, and 5), Glycoside Hydrolases (GH) family 5 and Auxiliary Activities (AA) family 1 showed a significant decreased in abundance at the organic horizon (Figure 2.5A). In contrast, the AA11 represent the only family above fold change threshold to increase with the experimental treatment in the organic horizon. At the mineral horizon, the

CAZy enzymes depicted an overall decrease in PL (i.e., families 4 and 14), AA (i.e., families 1 and 5), GH (i.e., families 5, 128, 72, and 28), and Carbohydrate Esterases (CE) family 8 (Figure 2.5B). In contrast, the carbohydrate binding module (CBM) group 6 increased in abundance at the mineral horizon.

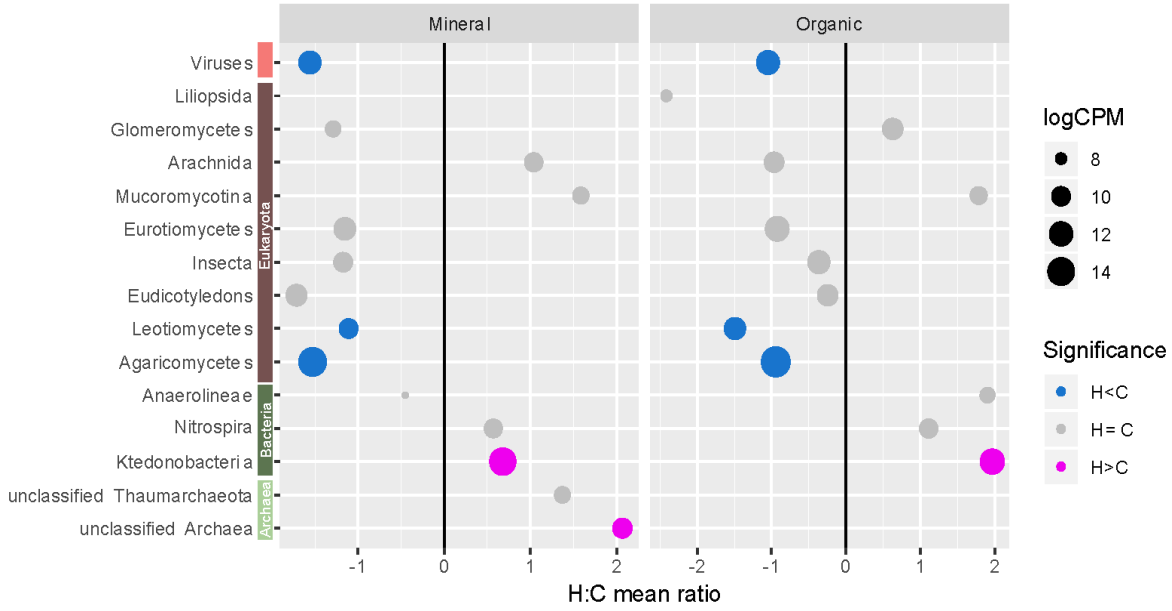

Figure 2.1: Changes in microbial biodiversity in response to the experimental treatment. The bubble plot depicts on the x-axis the H:C mean ratio, on the y-axis the taxonomic groups, bubbles (or circles) the abundance, and colored bubbles the statistical significance. The H:C ratio reflects the fold change between the heated plots relative to the controls (H:C ratio > 0 means increase in abundance with warming and H:C ratio <0 means otherwise). Bubbles captures both, the treatment significance (FDR≤0.1; red means increase, blue decrease, or gray no change in abundance with the warming effect) and microbial abundance in log2 count per millions (logCPM).

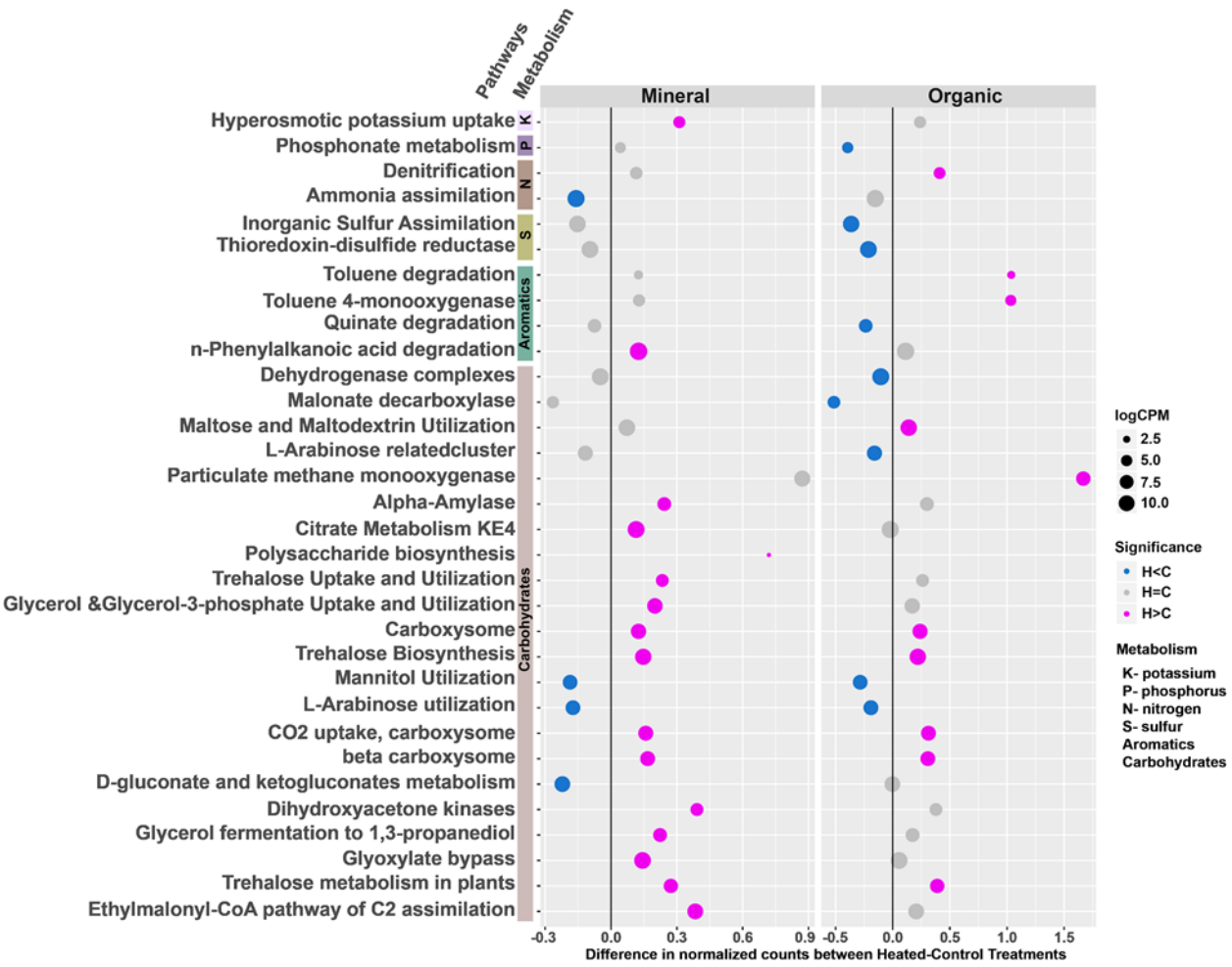

Figure 2.2: Shifts in microbial metabolism related to biogeochemical processes. Bubble plot for selected SEED metabolic pathways in both mineral and organic horizons. Pathways were selected based on SEED Subsystems associated to metabolism with significant (FDR≤0.1) warming effect. Plot shows the logarithmic fold change on the x-axis (similar to Figure 1) and pathways sorted by Subsystems (Potassium (K), Phosphorus (P), Nitrogen (N), Sulfur (S), Aromatics, and Carbohydrates) on the y-axis. Significance is capture on red and blue colored bubbles, while size the abundance.

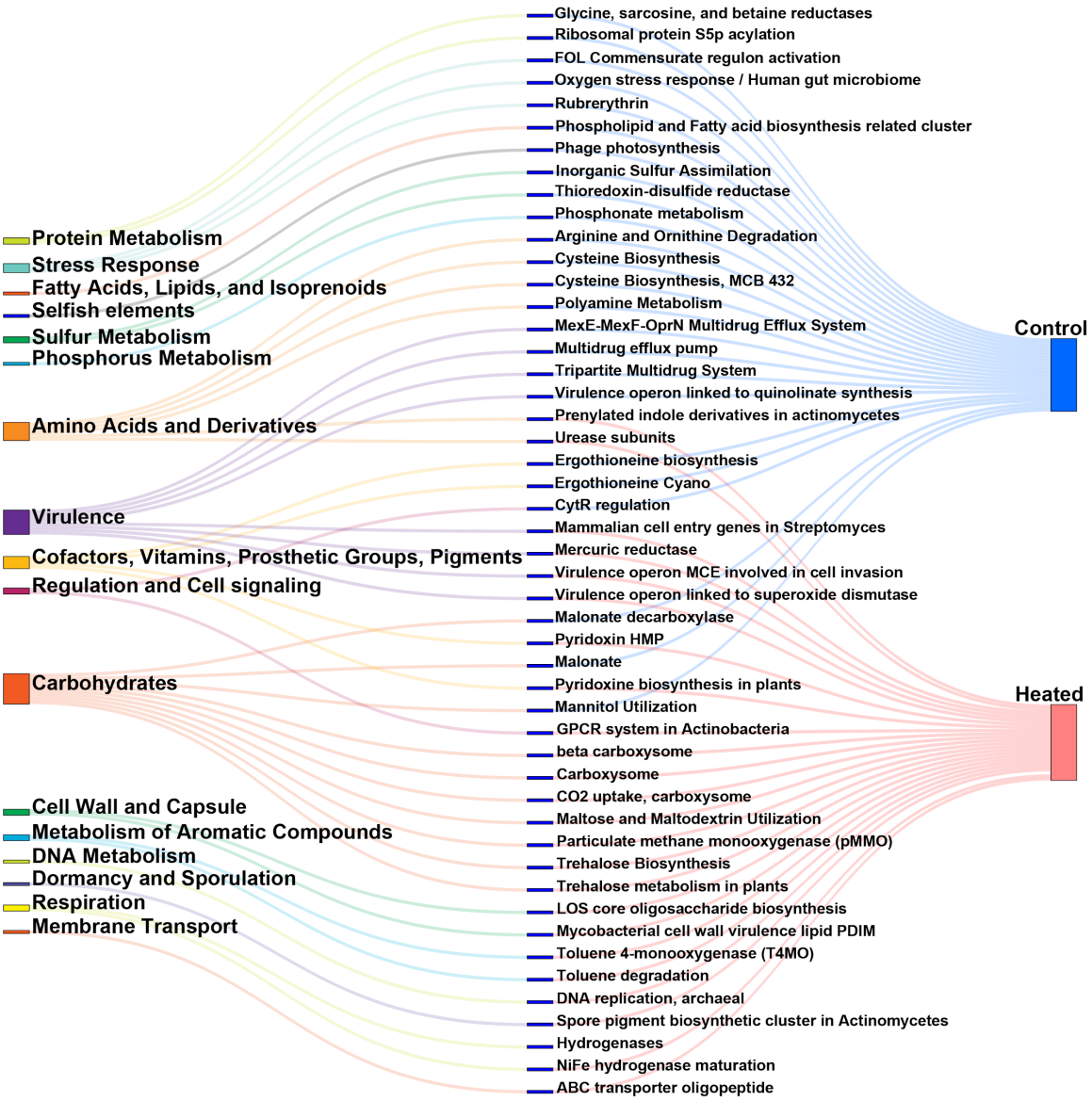

Figure 2.3: Sankey visualization of organic metatranscriptomes using the SEED functional annotation. From left to right, the SEED subsystems, SEED pathways, and treatment effect.

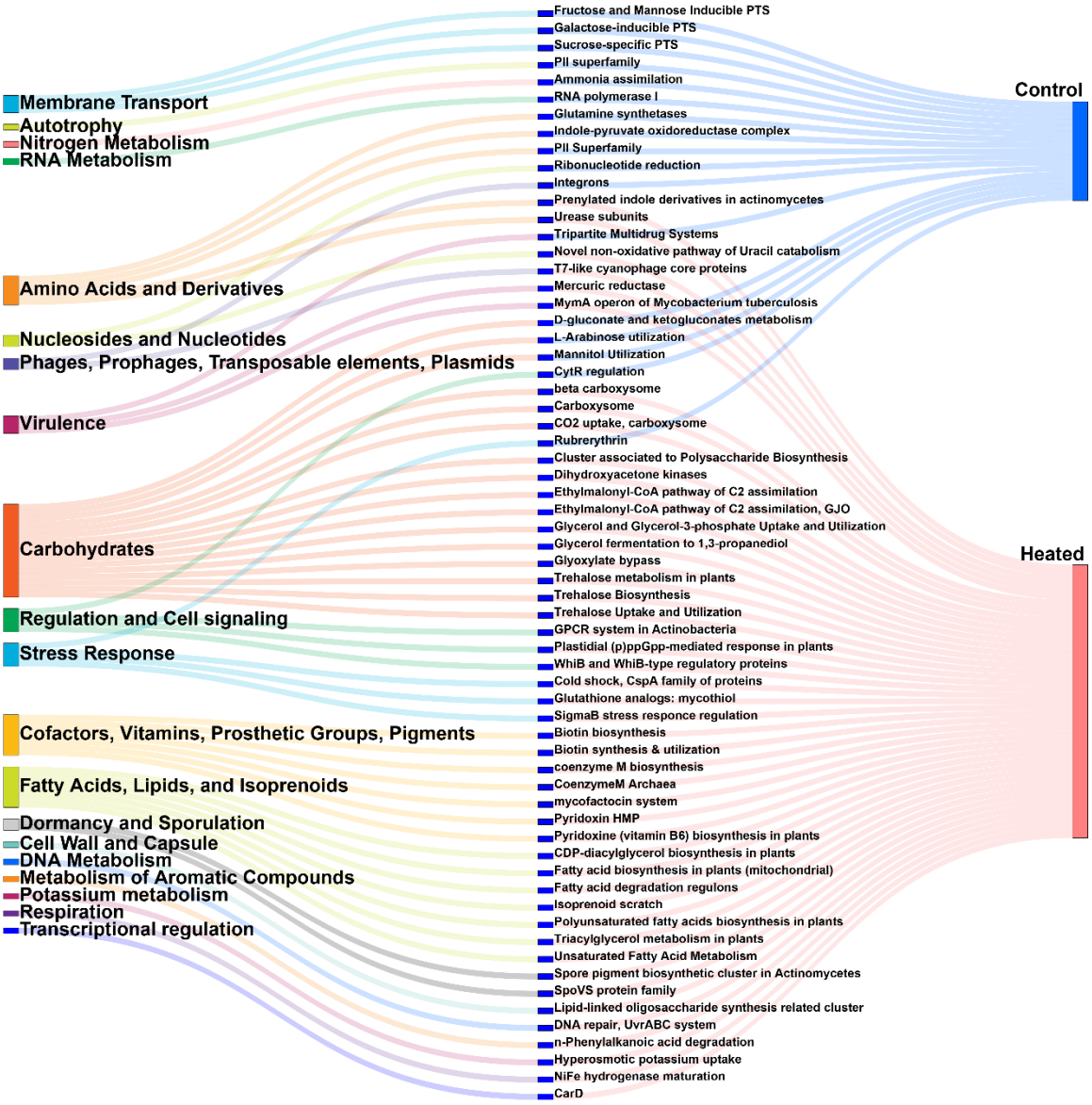

Figure 2.4: Sankey visualization of mineral metatranscriptomes using the SEED functional annotation. From left to right, the SEED subsystems, SEED pathways, and treatment effect.

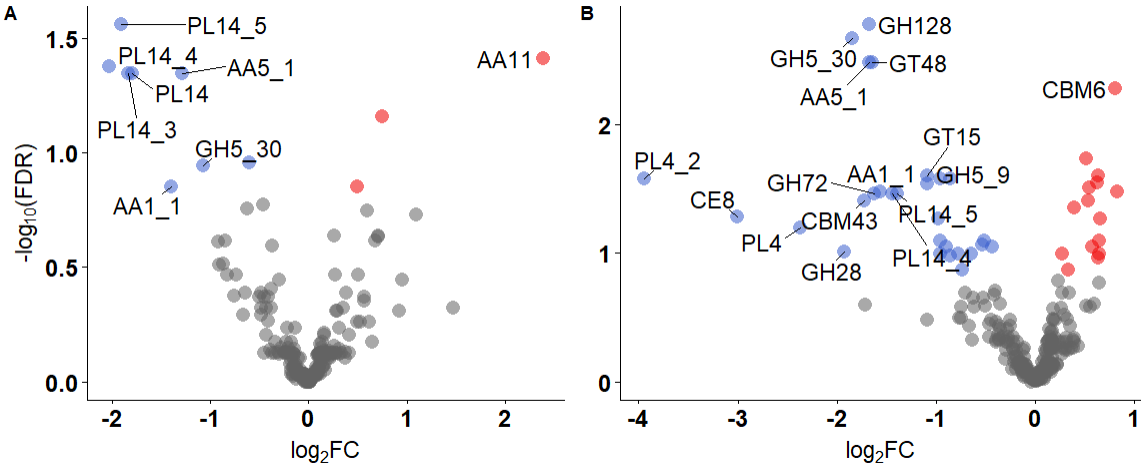

Figure 2.5: Warming effect on carbohydrates active enzyme families at the A) organic and B) mineral soil horizons.

2.4.2 Soil Chemistry

The overall number of detected carbon molecules was 16,814 (100%), including 9,723 (57.8%) with assigned molecular formula. Mass peaks outside the 200-900 m/z range were removed, resulting in 6,708 (39.9%) retained. The overall abundance of detected carbon compounds changed with the experimental treatment across soil horizons. In the organic horizon, the number of unique carbon molecules in the heated and control plot was 373 and 1,010, respectively (Figure 6A,C). In the mineral horizon, the number of unique carbon molecules in the heated and control was 1,016 and 368, respectively (Figure 6B,D). Abundance of detected soil carbon decreased 13% in the organic horizon with the warming treatment, while carbon molecules increased 17% in the mineral horizon with warming effect. From the unique carbon compounds (Figure 6C,D), the compounds were further expanded into their carbon categories breakdown. Carbon compounds assigned to lipids depicted a significant ($p \leq 0.1$) increase with the treatment effect in both soil layers. Amino sugars compounds decrease only in the organic horizon. In terms of element composition, the C compounds depicted significant changes across most categories among the unique carbon moieties. For

instance, the CHO and CHON categories reflected the most striking response to the treatment effect. The CHO index was calculated to inform about the oxidation state as reported on Mann et al. (Mann et al., 2015). This index reflected an increase in the number of reduced compounds with warming in both soil layers.

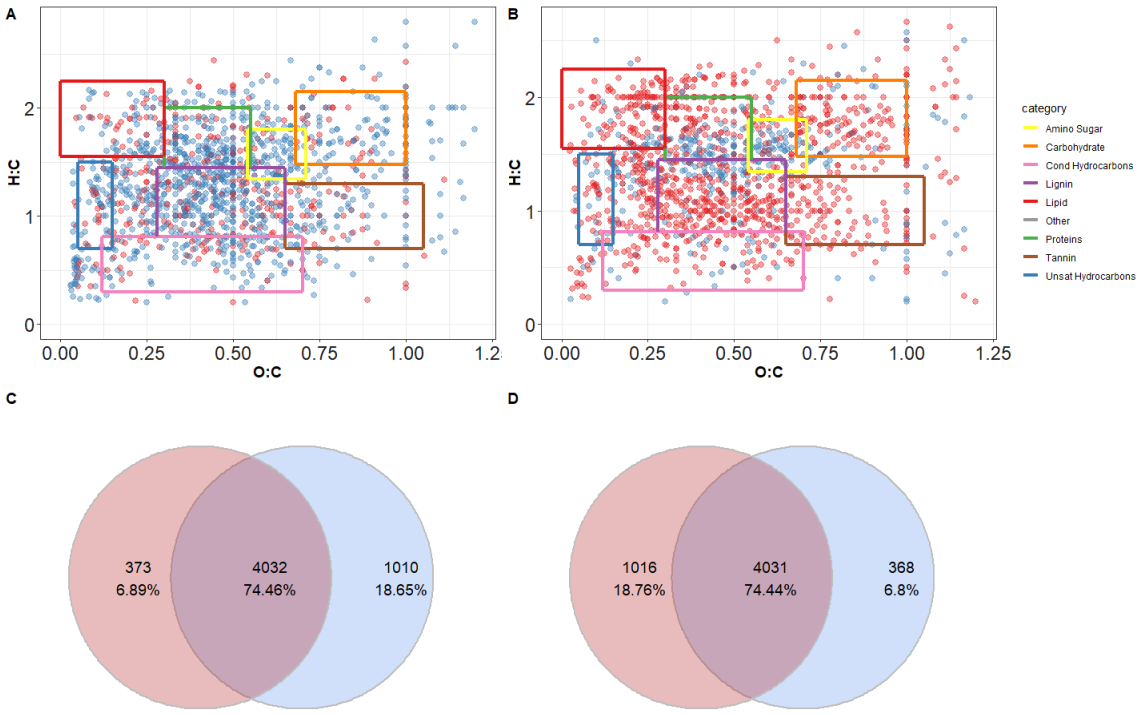

Figure 2.6: Changes in soil chemistry at the warming site. Van Krevelen diagrams for unique carbon compounds detected at the A) organic and B) mineral soil horizons. Compounds were colored by the treatment effect, red means only detected at the heated plots and blue means only detected at the control plots. Boxes within the Van Krevelen plots depict the carbon categories. Venn diagrams further summarize the relationship among compounds for the C) organic and D) mineral horizon.

2.4.3 Integration

Soil chemistry and microbial profiles were integrated using the Kyoto Encyclopedia of Genes and Genomes (KEGG) as metabolic framework (Okuda et al., 2008). The metabolic framework provides ecological information of microbial metabolism associated to enzymes and biochemical compounds. Linking the enzymes and compounds to ecosystem-level processes. For example,

the nitrite reductase (KEGG EC:1.7.2.1) enzyme catalyzed the chemical compounds nitric oxide (KEGG ID: CPD:C00533) to nitrite (KEGG ID: C00088), which play an important role in the nitrogen cycle (KEGG pathway: Nitrogen metabolism). Based on the FTICR-MS dataset, the biochemical compounds detected with assigned molecular formula were searched for hits against the KEGG compounds. In turn, the metatranscriptomes dataset informed about the enzymes for which the enzyme commission (EC) identifier was used as reference against the KEGG metabolic pathways. The manually curated SEED functional annotation was used as reference for enzymes. Since more than one enzyme can act on more than one compound and vice versa, our approach parsed chemical reactions to account for such interactions. Therefore, the visualized integrative network contains enzyme-nodes shared with multiple carbon compounds. When both FTICR-MS and metatranscriptomes datasets are taken together, this enable our integrative approach to inform chemical reactions using state-of-the-art high throughput technologies.

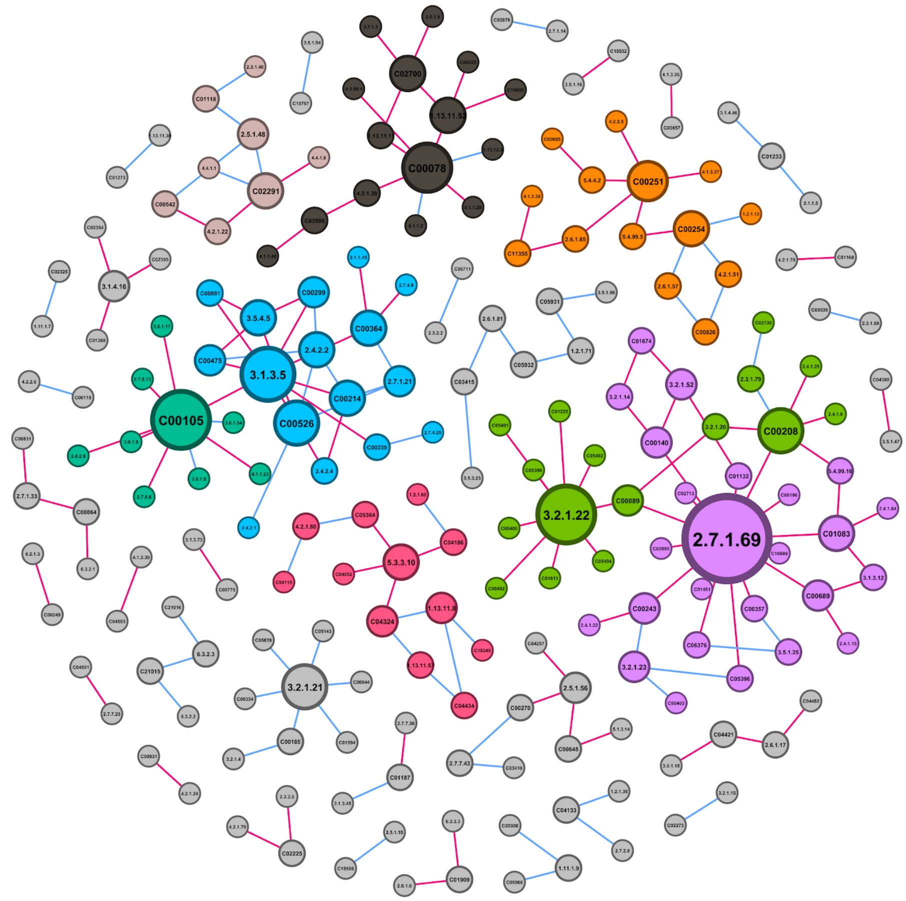

Figure 2.7: Integrative network with mass spectrometry and metatranscriptomic data. Nodes depict carbon compounds and enzymes captured with both FTICR-MS and metatranscriptomic in the organic horizon. Labels follow the KEGG nomenclature. Edges connect a compound to enzyme(s) and color capture the treatment effect based on enzyme activity. Enzyme-compound communities with at least 4% of nodes were highlighted as non-gray.

Table 2.1: Taxonomic changes in both soil horizons. Abundance of taxonomic groups in bold significantly (FDR≤0.1) change with the warming treatment and color indicate the treatment effect (red means group increase abundance at the heated plots and blue means group decrease in abundance at the heated plots).

Organic					
Domain	**Phylum**	**Class**	**Order**	**Family**	**Genus**
Bacteria (n.s.)	Acidobacteria (n.s.)	Acidobacteriia (n.s.)	Acidobacteriales (n.s.)	Acidobacteriaceae (n.s)	**Granulicella**
	Actinobacteria (n.s.)	Actinobacteria (n.s.)	**Corynebacteriales**	**Mycobacteriaceae**	**Mycobacterium**
	Chloroflexi	**Ktedonobacteria**	**Ktedonobacterales**	**Ktedonobacteraceae**	**Ktedonobacter**
	Proteobacteria (n.s.)	Alphaproteobacteria (n.s.)	Rhizobiales (n.s.)	**Beijerinckiaceae**	**Methylocapsa**
				Bradyrhizobiaceae (n.s.)	**Afipia**
		Betaproteobacteria (n.s.)	**Burkholderiales**	**Burkholderiaceae**	-
		Gammaproteobacteria (n.s.)	Pseudomonadales (n.s.)	**Moraxellaceae**	-
			Xanthomonadales	**Rhodanobacteraceae**	-
Eukaryota (n.s)	Ascomycota (n.s.)	**Leotiomycetes**	-	-	-
	Basidiomycota	**Agaricomycetes**	**Agaricales**	-	-
Viruses	-	-	-	-	-
Mineral					
Domain	**Phylum**	**Class**	**Order**	**Family**	**Genus**
Archaea (n.s.)	**Thaumarchaeota**	-	-	-	
Bacteria (n.s.)	Actinobacteria (n.s)	**Actinobacteria**	**Corynebacteriales**	Mycobacteriaceae (n.s.)	**Mycobacterium**
		Thermoleophilia	-	-	-
	Chloroflexi (n.s)	**Ktedonobacteria**	-	-	-
Eukaryota (n.s.)	**Ascomycota**	**Leotiomycetes**	-	-	-
		Sordariomycetes		-	-
	Basidiomycota	**Agaricomycetes**	**Agaricales**	-	-
Viruses	-	-	-	-	-

Table 2.2: IMG accession numbers for soil metatranscriptomes.

Sample ID	IMG accession	Sample ID	IMG accession
BWC12M	3300022722	BWH11M	3300022717
BWC12O	3300022505	BWH11O	3300022522
BWC14M	3300022510	BWH17M	3300022724
BWC14O	3300022499	BWH17O	3300022523
BWC19M	3300022508	BWH26M	3300022506
BWC19O	3300022502	BWH26O	3300022500
BWC27M	3300022507	BWH28M	3300022531
BWC27O	3300022509	BWH28O	3300022511
BWC30M	3300022726	BWH2M	3300022504
BWC30O	3300022530	BWH2O	3300022503
BWC4M	3300022532	BWH32M	3300022712
BWC4O	3300022721	BWH32O	3300022498
BWC7M	3300022533	BWH4M	3300022525
BWC7O	3300022501	BWH4O	3300022527

2.5 Discussion

At the Barre Woods site, the long-term warming effect has resulted positive feedbacks to woody tissue of plants and release of carbon into the atmosphere (Melillo et al., 2011). This carbon release has been associated to labile soil organic matter. In this study we demonstrated that long-term warming resulted in significant changes in soil biodiversity and soil organic matter decay in organic horizon. Furthermore, the FTICR-MS and metatranscriptomes were successfully integrated using KEGG as a metabolic framework. This allowed us to link microbial metabolism and soil chemistry in addition to interrogate each data set individually. Based on the soil chemistry profiles, the carbon compounds were depleted by 13%, whereas the carbon compounds increased by 17% in the mineral horizon. This carbon turnover suggests a translocation based on microbial transformations in the upper (organic) horizon to the lower (mineral) horizon. Although abiotic factors can be associated to the carbon shift, the detection of enzymes acting on those carbon compounds provide an extra layer of confidence about the biotic transformations.

The soil chemistry analysis detected carbon moieties commonly found in the forest soil. For instance, the amino sugar compounds can be found in the cell wall of plant, bacteria and fungi, which upon senescence serve as a nitrogen source to microbes (Hu et al., 2018). At the Harvard Forest site, we measured a decrease in amino sugar, which the KEGG framework linked phytochemical compounds. The lipid compounds were linked to oxidoreductase reactions with NAD+ and NADP+ as acceptor. These cofactors play an essential role in redox metabolism and electron transport, therefore, cellular respiration (Rich, 2003). Consonant with previous observations, the long-term warming effect has resulted in the increase of cellular respiration at the field site (Melillo et al., 2017).

The integrative approach provided a new framework to synthesize soil chemistry and metatranscriptomes. This captured complete chemical reactions, i.e. compounds and enzymes, from a complex mixture of organic material and metabolites. From the integrative network, the compound-enzyme communities depicted functions necessary for cellular viability and stability processes, including amino acid synthesis and carbon metabolism (i.e., transport and degradation). This involved enzymes like chitobiose and chitinase employed by microbes to decompose soil organic carbon. As long-term warming favored more recalcitrant compounds, the observed increase in abundance for these enzymes were expected at the field site (Pold et al. 2017).

Comparative analysis of metatranscriptomes detected significant shifts in abundance of fungi, prokaryotes, and viruses. The decrease in fungi involved groups with a wide range of ecological functions, including ectomycorrhizal parasites, symbionts, and saprobes (Hibbett, 2006; Wang et al., 2006). At the Harvard Forest, the fungal decrease has been reported to reduce carbohydrate degradative enzymes with the temperature effect in the organic horizon (Pold et al., 2016). In terms of viruses, the decrease in abundance was linked to positive-sense single stranded RNA (ssRNA) viruses. Members of ssRNA viruses has been described as symbionts of arthropods and vertebrates (Jose et al., 2009). Recently, our collaborators provided in depth insights of the viral diversity in the forest soil, which can harbor giant viruses (Schulz et al., 2018). In previous studies viruses has been described to influence the carbon metabolism by association to bacteria and auxiliary CAZy groups (Jin et al., 2019).

The Acidobacteria, Actinobacteria, and Proteobacteria were the three major bacterial phyla playing a key role in metabolic pathways with significant treatment effect. Changes associated to Bradyrhizobiales and Beijerinckiaceae were linked to ammonia assimilation and

denitrification processes. Previous studies described both bacterial groups as nitrogen fixers and methanotrophs (Jayasinghearachchi and Seneviratne, 2004; Lau et al., 2007). Abundance of pathogenic bacteria increase with the experimental treatment. This included Mycobacterium and Thermoleophilia groups. Mycobacterium members are well known pathogenic bacteria in human with the ability to degrade recalcitrant carbon in the gut (Gielnik et al., 2019). However, there are known members with similar characteristics isolated from soils (Wolinsky and Rynearson, 2015). In turn, the Thermoleophilia group has been reported to correlate with availability of dissolved organic matter and phosphorus limitation (Cui et al., 2018; Wu et al., 2018). Although not pathogenic, the Ktedonobacteria is another soil-dwelling group that significantly increase in abundance and whose functional features (i.e. spore formation) allowed them to thrive under unfavorable conditions (Chang et al., 2011). In fact, the differentially abundant pathways associated to spore formation were linked to this group within Chloroflexi. Further, this group is capable to fix carbon monoxide (King and King, 2014). In fact, the taxonomic affiliation of the most abundant gene with significant treatment, RubisCO, was associated to this bacterial group.

The integrative approach presented here provides a new framework to link cellular and chemical information from soil. Furthermore, this framework can be easily extended into other systems. As new high throughput technologies continue to evolve, more efforts are required to synthesize the large amount of information. Although challenging, there are plenty of space to further integrate multi-omics data. The integration of large-scale genomics data sets will provide stronger insights in natural communities and advance research in microbial ecology.

CHAPTER 3

TIME SERIES OF SOIL METATRANSCRIPTOMES: INSIGHTS OF MICROBIAL COMMUNITIES EXPOSED TO LONG-TERM WARMING AT THE PROSPECT HILL SITE

3.1 Abstract

Ecosystem dynamics are influenced by abiotic and biotic factors. The main abiotic factors involve temperature and moisture. Both abiotic factors can control forest growth above ground and reshape the microbial community structure below ground. These complex dynamics conspire to shift nutrients (e.g., soil organic carbon) and enable new communities to establish. Here, we analyzed a time series of metatranscriptomics to determine changes in soil microbial communities in a natural environment and after long-term warming exposure. This leverage on the Prospect Hill warming site located in a mid-latitude deciduous forest. The samples were collected from six time points and one treatment with replication. The treatment effect consisted of an increase of 5℃ in the heated plots relative to the ambient temperature measured in the controls. Results depicted a similar community structure across the seasonal time points. However, the warming effect resulted in a strong differentiation of soil metatranscriptomes. Significant community changes occurred within the Proteobacteria phyla. Warming increased the abundance of Alpha-proteobacteria and decrease of Beta-proteobacteria groups. This time series provided new insights in microbial structure after long-term warming exposure and seasonal variation.

3.2 Introduction

Soil microbes are an essential component in mediating biogeochemical processes. These microorganisms respond to soil carbon availability, carbon chemistry, and nutrient content (Classen et al., 2015; Melillo et al., 2017). A single gram of soil is estimated to harbor up to 1×10^9 microbial species. This large biodiversity has been link to major roles in ecosystem functioning

(Maron et al., 2018). This microbial role involves nutrient acquisition (Heijden et al., 2008), carbon cycling (Bardgett et al., 2008), phosphorus cycling (Correa et al., 2016), and nitrogen cycling (Hayatsu et al., 2008). Moreover, the abiotic factors exert control over microbial dynamics to move nutrients and eventually ecosystem-level processes.

There is seasonal variation influences several abiotic factors in terrestrial ecosystems. This include moisture, temperature, pH, and vegetation at Harvard Forest. However, temperature and moisture are by far the main abiotic mechanisms, which in turn control pH and vegetation. Several studies indicate that both abiotic factors regulate carbon mineralization (Benbi et al., 2014; Curtin et al., 2012; Taggart et al., 2012). Besides the regulation of soil organic matter, moisture itself can constraint the mobility of bacteria and in excess can destroy soil aggregates (Viswanath and Pillai, 1972; Yang et al., 2017). The soil aggregates serve as pocket of nutrients and niche for microbial communities (Upton et al., 2019). Therefore, the temporal variation shapes the biodiversity of both above and below ground.

Insights of natural communities in soil continue providing a new dimension of the high biodiversity below ground (Jing et al., 2015; Peters et al., 2019; Schulz et al., 2018). More recently, a new framework to categorize the large amount of genomic information was proposed (Fierer et al., 2007). The ecological concept presented by Fierer et al. provided a new classification framework to soil-dwelling bacteria, which can be categorized as copiotrophs or oligotrophs. This framework was adapted from the r- and K- selection categories established for macroecology (Pianka, 1970). The oligotrophs and copiotrophs differentiate by their physiological traits. Oligotrophic microorganisms are capable to survive under poor nutritional environments, whereas the copiotrophic cannot proliferate due to high nutritional requirements.

Microbial metabolism and biodiversity are affected by temperature (Okie et al., 2015). In the tundra, the decomposition rate of soil organic carbon is limited by low temperatures (Koyama et al., 2014). In contrast, soil organic carbon and other nutrients are limited to soil microorganisms in the tropics (Kaspari et al., 2008). Here, we used a metatranscriptomics approach to capture changes in the microbial community structure across a time series that captured the seasonal variation at the Prospect Hill warming site. The Prospect Hill warming site was established in 1991 and form part of the Long-Term Ecological Network sites (Melillo, 2002). This outdoor experiment consists of increasing the soil temperature 5℃ above the ambient temperature. As result, we can further explore the microbial community structure in response to both a seasonal variation and long-term warming exposure. While the vast majority of soil microbes remains uncharacterized, new molecular techniques enabled the study of natural communities (Zhou et al., 2015). To this end, we profiled the 16S and 18S regions obtained from the soil metatranscriptomes. Our results suggest that long-term warming produce a stronger response to the community structure relative to the time series. Also, we provided suggestions to better capture the microbial activity and biodiversity using metatranscriptomics.

3.3 Methods

3.3.1 Soil Samples

Harvard Forest is home of the Long-Term Ecological Research (LTER) site used throughout these experiments. Soil samples were collected for 6 time points at the Prospect Hill warming site (Table 3.1). The time points capture the temporal changes at the Harvard Forest, which is a deciduous forest located in Petersham, Massachusetts. Also, the samples were collected from the disturbed control plots and heated plots. Here we referred to the disturbed control plots as controls. It is important to make this distinction as samples were collected from

plots with installed coils like in the heated plots, but these coils were never turned on. At the time of soil collection, the heated plots experienced the increase in soil temperature, 5℃ above the ambient temperature, for the past 23 years. The soil cores were divided by horizon, with the organic horizon and mineral separated by visual inspection. This method has shown to work well as both soil layers depict very distinct colors. On site, the samples were sieved to remove rocks and placed in a cooling bath. This bath consisted of a mix of ethanol and dry ice. Then, the samples were transported to UMass and storage at -80 ℃ until nucleic extractions.

Table 3.1: Dates of sample collection at the Prospect Hill warming site. The columns show the time point order, date, and week of the year.

Time Point	Date of sampling	Week number
T1	04/28/14	18
T2	06/03/14	23
T3	06/30/14	27
T4	08/25/14	35
T5	09/22/14	39
T6	10/27/14	44

3.3.2 RNA Extractions and Sequencing

Total RNA was extracted from 144 samples using the RNeasy PowerSoil RNA extraction kit (QIAGEN). Also, an artificial phage was added to the complementary DNA (cDNA) libraries to measure the RNA extraction efficiency post-sequencing. The phage was added to account for 0.05% of the RNA yield per sample. Then, the cDNA libraries were assessed for quality using a bioanalyzer and qPCR machine. This process selected for samples that meet or exceed a minimum concentration of 1.8 pM per libraries. The Illumina NextSeq platform and 500/550 Mid Output kit with 300 cycles was used to sequence the libraries. This resulted in 57 organic samples sequenced (Table 3.2).

Table 3.2: Breakdown of sequenced soil samples per horizon and treatment at the Prospect Hill time series. The columns show the time point, total number of sequenced samples by time point, and number of sequenced samples by treatment.

Time Point	Samples per time point	Samples per treatment
T1	10	Control - 7 Heated - 3
T2	8	Control - 4 Heated - 4
T3	12	Control - 6 Heated - 6
T4	9	Control - 4 Heated - 5
T5	10	Control - 6 Heated - 4
T6	8	Control - 4 Heated - 4

3.3.3 Bioinformatics and Statistics

Raw pair-ended sequences were checked for quality using the FastQC tool (Andrews, 2010). Then, the artificial phage reads were removed by searching the sequenced samples against the phage-target sequence with the BBduk tool within the BBTools suite (Bushnell, 2016). The phage was added only to determine RNA yield efficiency. GNU parallel library was employed to optimize the analysis process in multiple samples (Tange, 2011). BBduk suite was employed again to remove adapters and trim low-quality ends (10bp). Then, the FLASH tool was used to concatenate both pair-ended sequences. SortmeRNA tool served to identify both ribosomal RNA (rRNA) and non-rRNA sequences (Kopylova et al., 2012). Finally, the RiboTagger tool was used to align the cleaned sequences against the V4 variable region (Xie et al., 2016). The RiboTagger algorithm provides a higher level of sensitivity over RDP or SILVA algorithms. Furthermore, this tool can target taxonomic groups across three domains (i.e., bacteria, archaea, and eukaryotes). Taxonomic resolution is comparable to assignment of operational taxonomic units (OTUs). RiboTagger integrates well with MEGAN to easily generate the community matrices by taxonomic rank and export the taxonomic frequencies (Huson et al.,

2016). The R statistical language was employed to determine significant taxonomic groups and visualize the results for the time series (R Core Team and R Development Core Team, 2008).

3.4 Results

Soil temperature changed across the time series (Figure 3.1). The lowest temperature was recorded at T1 (wk-18) and T6 (wk-44), which represent each end of the time points. In turn, the warmest soil temperature was detected at T3 (wk-27) and T4 (wk-35) for which samples were collected during the summer season. Furthermore, the soil water content decrease at T3 and T4 in respect to the other time points. In contrast, the soil water content increased at the shoulder time points, i.e. T1 and T6, as rain events occur more often during the months of soil collection (Table 3.1). Therefore, the time series captured the seasonal variation at the Harvard Forest.

Sequencing resulted in 57 samples with 24,271,193 sequences across the time series (Table 3.3). The ribosomal RNA (rRNA) sequences were assigned to 846 unique taxonomic groups (Magoč and Salzberg, 2011). We were able to detect taxonomic groups associated to bacteria and eukaryotes (Figure 3.2). Although ordination techniques reflected soil samples not clustering by the temporal effect, the soil metatranscriptomes depicted a strong warming effect across the time series (Figure 3.3). When samples were split by time point, the pattern was clear and more obvious (Figure 3.3B).

Warming altered the diversity of taxonomic classes within time points (Figure 3.4). The diversity metrics reflected an overall increase in abundance with the temperature treatment. However, the warming effect resulted in a significant (p value<0.05) warming effect only in T1 and T6. We further tested the temperature effect within taxonomic groups. At the class rank, we detected several groups commonly found in soil (Figure 3.5). Changes in abundance were

detected among this taxonomic groups. For instance, the abundance of Actinobacteria increased with warming at T1, whereas the Betaproteobacteria and Acidobacteriia groups decreased in abundance. At T2, the abundance of Planctomycetes and Spartobacteria groups increased, while the Agaricomycetes decreased in abundance. In T3, the Spartobacteria class was the only taxonomic group with significant treatment effect, which slightly increased in abundance with warming. At T4, the Bacilli and Alpha-proteobacteria groups increased in abundance, as opposed to the Betaproteobacteria. In T5, the Bacilli class was the only significant group whose abundance increase with respect to warming. At T6, the Plantomycetia increased in abundance with the treatment effect, while the Agaricomycetes and Sphingobacteria classes decreased in abundance with warming.

Taxonomic abundance of metatranscriptomes were linked to soil-dwelling bacteria and fungi. Highly abundant groups included Bradyrhizobium (alpha-proteobacteria), Xanthobacter (Alphaproteobacteria), Sorangium (Deltaproteobacteria), Burkholderia (Betaproteobacteria), and Russula (Ascomycetes) (Figure 3.5). Furthermore, the abundance of these groups tended to increase with the temperature effect within time points and in some instances following a pattern across the time points (Figure 3.6). For individual time points, the warming effect resulted in significant differences for Bradyrhizobium, Burkholderia, and Xanthobacter (Table 3.4). However, other general trends arise across the time series. For instance, the abundance of the Bradyrhizobium genus was constantly higher in the heated plots relative to the control plots. In the heated plots, this Alphaproteobacteria group showed a consistent increase from wk-18 to wk-35, then the abundance decreased in wk-39 close to the abundance in wk-18 and gradually started to increase as depicted in wk-44 (Figure 3.6). Relative abundance of the fungal genus Russula depicted a bell-like shape in the control plots, with peaks in wk-27 and wk-35. Similar to previous studies, the fungi markers decrease in abundance with the temperature effect.

Detection of Russula in the heated plots fluctuated across the temporal factor. Sorangium was another taxonomic group for which taxonomic distribution resembles a bell curve with higher abundance at wk-27 and wk-35. In contrast, the abundance of Sorangium gradually decrease in the heated plots from wk-18 until wk-39 and abundance increase once again in wk-44.

Table 3.3: Sequence breakdown and metadata associated to the time series. The columns depict the sample identifier (Sample ID), unprocessed sequences (Raw Reads), artificial phage added to libraries (Phage Reads), concatenated and cleaned pair-ended reads (Merged Reads), GC percentage (GC%), ribosomal RNA sequences as detected with sortmeRNA (rRNA Reads), Time Point, Treatment , soil weight in grams used for extraction (Soil), and RNA concentration per sample (RNA).

Sample ID	Raw Reads	Phage Reads	Merged Reads	GC %	rRNA Reads	Time Point	Treatment	Soil (g)	RNA (ug)
PH-T1-DC10-O	27478506	72309	12955144	53	12168353	T1	Control	1.18	26.55
PH-T1-DC13-O	52973222	163363	24065314	55	22218052	T1	Control	1.31	7.527
PH-T1-DC17-O	1986158	12824	910182	50	842393	T1	Control	1.42	1.563
PH-T1-DC3-O	38228302	94928	16388197	54	15382325	T1	Control	1.25	27.15
PH-T1-DC3-O	267410998	99202	114412482	54	107238528	T1	Control	1.25	27.15
PH-T1-DC5-O	11593464	42641	5259245	53	4918348	T1	Control	1.19	19.125
PH-T1-DC9-O	22071550	57486	9858470	53	9243819	T1	Control	1.43	17.475
PH-T1-H12-O	15393036	47941	7072116	54	6640795	T1	Heated	1.08	21.375
PH-T1-H16-O	90620464	397857	38728400	56	33426392	T1	Heated	1.64	5.895
PH-T1-H6-O	22764676	91261	10667447	54	10000417	T1	Heated	1.89	11.115
PH-T2-DC13-O	169444040	74970	79866460	53	69569423	T2	Control	1.17	11.526
PH-T2-DC3-O	348144656	44584	157490023	52	149106041	T2	Control	1.06	39
PH-T2-DC5-O	192207754	75388	88873066	53	82249246	T2	Control	1.31	21.225
PH-T2-DC9-O	207754124	7296	91077669	54	84165090	T2	Control	1.13	29.925
PH-T2-H1-O	243363274	76854	104136737	55	95514339	T2	Heated	1.08	17.925
PH-T2-H16-O	238137640	55726	105412899	54	100251257	T2	Heated	1.11	22.335
PH-T2-H6-O	244890706	82358	112429979	54	103300103	T2	Heated	1.18	28.125
PH-T2-H8-O	221499300	55480	104957793	54	97567132	T2	Heated	1.19	23.775
PH-T3-DC10-O	12571376	51535	5835469	54	5467110	T3	Control	1.18	17.385
PH-T3-DC13-O	14530798	39490	6683446	56	6238048	T3	Control	1.3	17.58
PH-T3-DC17-O	5497428	95943	2535128	58	1946597	T3	Control	1.28	18.45
PH-T3-DC3-O	12093250	30308	5484958	54	5109452	T3	Control	1.1	24.525
PH-T3-DC5-O	65423990	163067	28237720	52	25718484	T3	Control	1.17	48.375
PH-T3-DC9-O	18261058	80994	8384224	54	7735795	T3	Control	1.21	18.165
PH-T3-H1-O	11248368	31653	5015560	54	4634314	T3	Heated	1.17	28.575
PH-T3-H12-O	25215176	65862	12066855	53	11140194	T3	Heated	1.12	26.775
PH-T3-H15-O	6337246	19606	2954899	54	2785871	T3	Heated	1.18	31.725
PH-T3-H16-O	29712352	81941	14117306	54	13063592	T3	Heated	1.29	15.54
PH-T3-H6-O	29425968	94754	12980858	54	12097194	T3	Heated	1.06	23.925
PH-T3-H8-O	2294398	7171	1033392	54	970948	T3	Heated	1.05	43.35
PH-T4-DC17-O	11129646	60280	5101212	56	4245537	T4	Control	1.29	4.95
PH-T4-DC3-O	17376602	42575	8210472	54	7664231	T4	Control	1.05	21.375
PH-T4-DC5-O	66077288	188809	28410595	54	26110198	T4	Control	1.4	8.718
PH-T4-DC9-O	61719553	52996	9281257	52	8526557	T4	Control	1.17	14.97
PH-T4-H1-O	37204872	110698	16427013	54	15170337	T4	Heated	1.2	25.65
PH-T4-H12-O	459650	3943	194081	56	172100	T4	Heated	1.25	8.94
PH-T4-H16-O	8675206	25929	3965515	55	3707639	T4	Heated	1.195	13.83
PH-T4-H6-O	1077978	3156	507958	55	467663	T4	Heated	1.18	24.225
PH-T4-H8-O	18703476	57312	8606961	55	8082963	T4	Heated	1.25	36.3
PH-T5-DC10-O	7739300	42229	3636988	55	3084879	T5	Control	1.52	10.6515
PH-T5-DC13-O	62613750	206389	28624578	53	26743228	T5	Control	1.07	50.1
PH-T5-DC17-O	6030690	23606	2500579	56	2194071	T5	Control	1.36	14.9265
PH-T5-DC3-O	1944428	5113	800999	54	733124	T5	Control	1.42	11.3385
PH-T5-DC5-O	80373270	239239	33548673	52	30208920	T5	Control	1.2	22.95
PH-T5-DC9-O	68928160	180240	32653877	54	30058519	T5	Control	1.08	34.575
PH-T5-H1-O	18852826	64268	8442940	55	7747152	T5	Heated	1.49	24.675
PH-T5-H12-O	36433158	81436	16448292	53	15462647	T5	Heated	1.27	23.85
PH-T5-H16-O	53266360	131707	24950502	54	23590176	T5	Heated	1.22	33.15
PH-T5-H8-O	30232713	67225	7543142	54	6922551	T5	Heated	1.39	19.725
PH-T6-DC10-O	238326374	64990	113358929	54	106429755	T6	Control	1.12	58.575
PH-T6-DC13-O	98480806	15480	43663580	53	40419480	T6	Control	1.1	50.85
PH-T6-DC5-O	387798190	51202	179276720	53	166669534	T6	Control	1.1	17.37
PH-T6-DC9-O	262495378	127052	122144339	51	112146763	T6	Control	1.25	20.445
PH-T6-H1-O	148706530	26646	69791495	54	64166191	T6	Heated	1.07	30.3
PH-T6-H12-O	226352424	26408	102765550	54	95482043	T6	Heated	1	26.1
PH-T6-H16-O	176497716	3686	82173987	57	72607167	T6	Heated	1.46	7.92075
PH-T6-H8-O	227673250	62886	103286693	54	96264568	T6	Heated	1.2	27.18

Figure 3.1: Soil temperature measured in the control and heated plots at the Prospect Hill site. The plot depicts the average temperature (℃) on the y-axis and sample collection time point with week number within parenthesis on the x-axis.

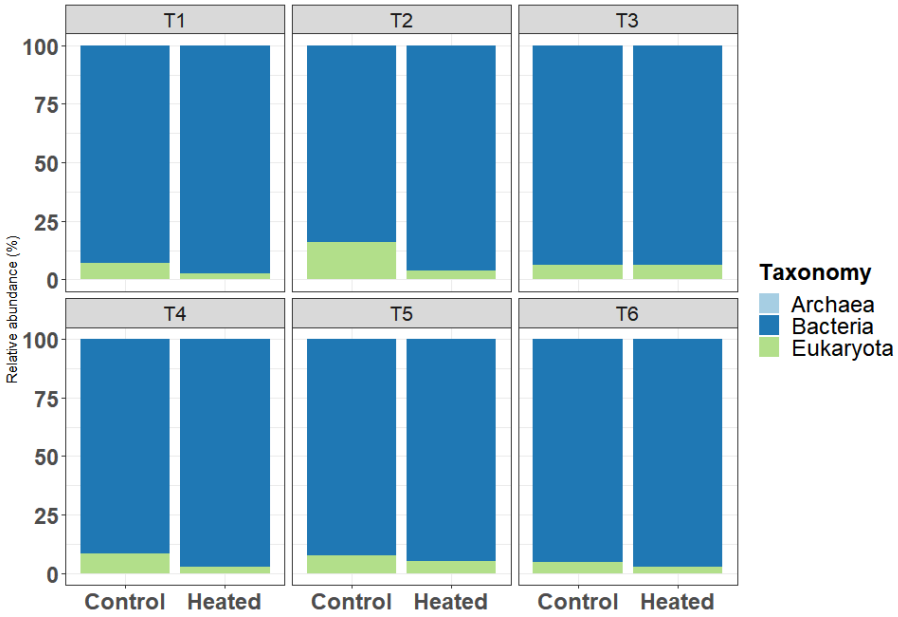

Figure 3.2: Abundance of taxonomic groups at the domain rank for each treatment across the time series. The stacked bar plot shows the relative abundance of transcripts in the y-axis and the treatment effect on x-axis for each time point (T1, week 18; T2, 23 wk; T3, wk 27; T4, wk 35; T5, wk39; and T6, wk 44).

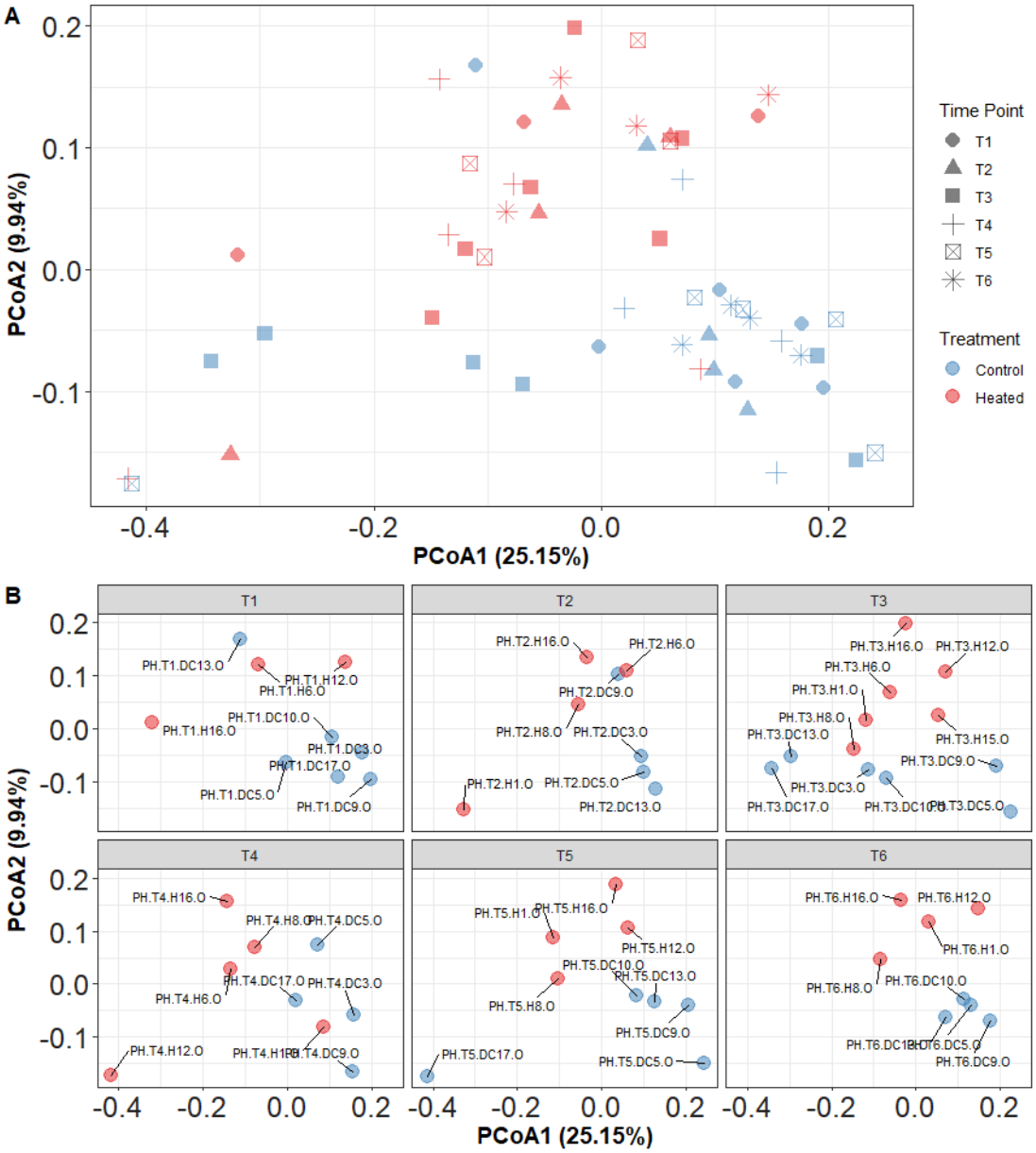

Figure 3.3: Microbial structure differ by the treatment effect across the time series. PCoA ordination plots were presented with A) all time points combined and B) splitted by time point. The Bray Curtis distance was used to determine similarities among soil metatranscriptomes.

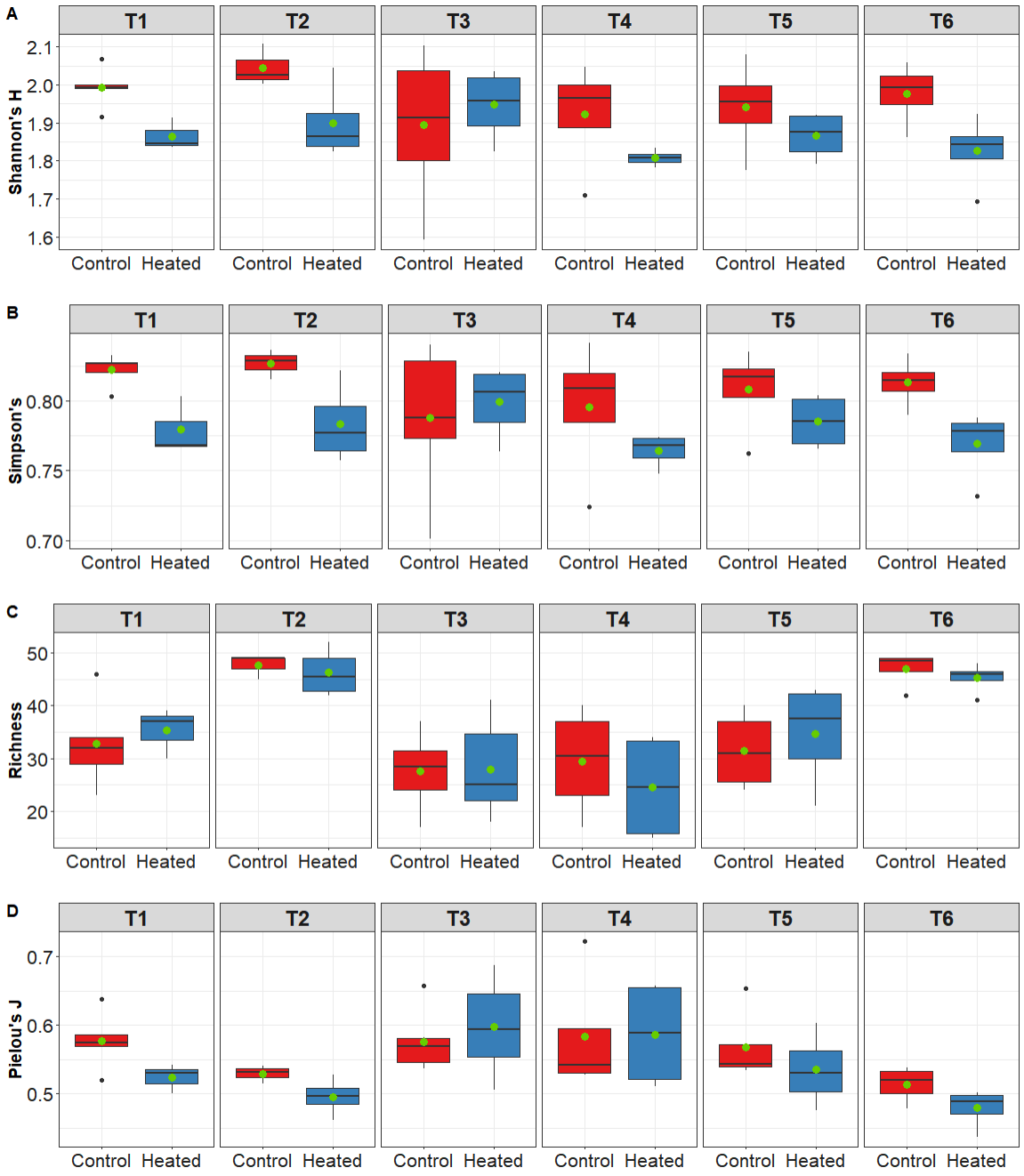

Figure 3.4: Diversity metrics for taxonomic groups at the class resolution. The metrics includes A) Shannon's, B) Simpson's, C) Richness, and D) Pielou's evenness indices for each time point.

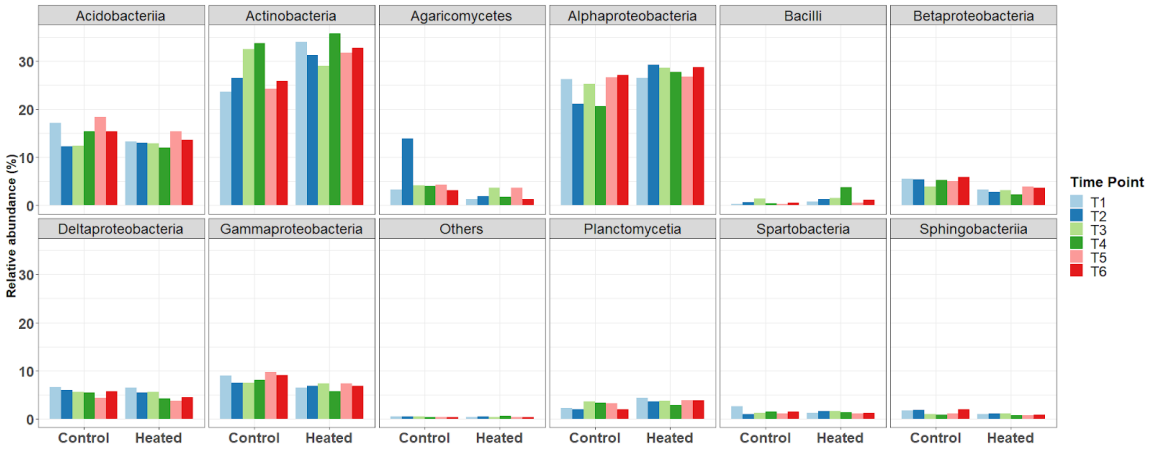

Figure 3.5: Microbial community shifted in abundance with the warming effect. Abundance of taxonomic groups at the class resolution. Groups with abundance below 1% were binned as others.

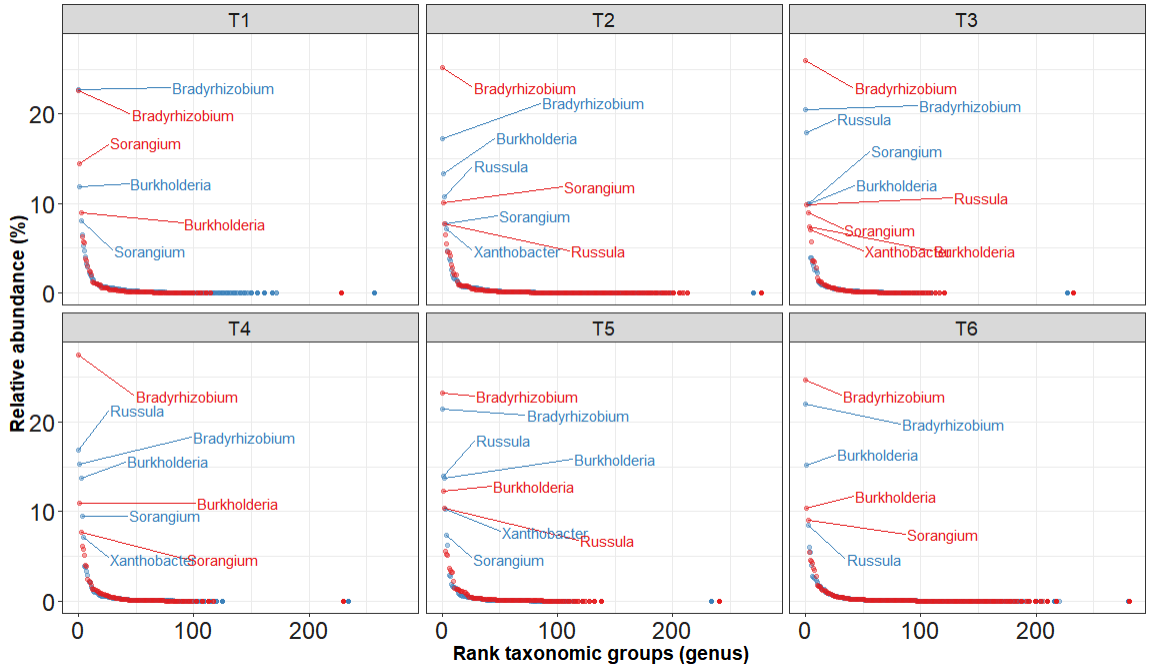

Figure 3.6: Rank abundance curve of taxonomic group annotated at the Prospect Hill site. Taxonomic groups at the genus resolution and labels added to taxa with relative abundance >7%. Treatment effect captured with the highlighted colors, red means annotated at the heated plots and blue means annotated in the control plots.

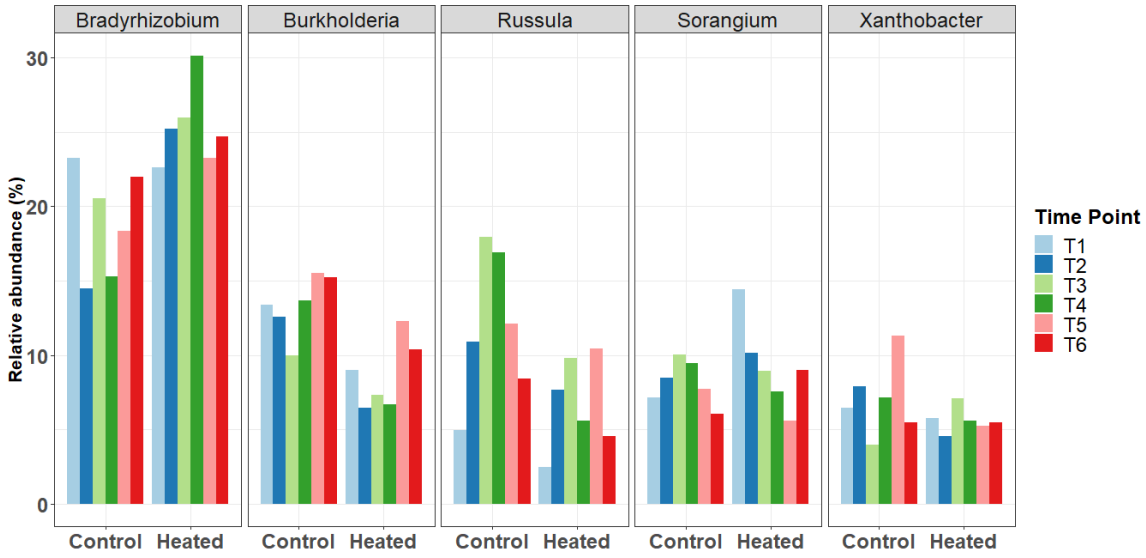

Figure 3.7: Spatiotemporal and treatment effect within the most abundant microbial groups. Taxonomic groups showed at the genus rank. The seasonal effect is captured with the highlighted colors and treatment effect showed within facet for each of the top 5 genera.

Table 3.4: Warming effect for core taxonomic groups exposed to the long-term warming treatment across the time series. The columns show the time points, core taxonomic group at the genus rank, and p value based on a Wilcoxon test.

Time Point	Taxonomy	Pvalue
T4	Bradyrhizobium	0.028571429
T4	Burkholderia	0.028571429
T1	Burkholderia	0.035714286
T2	Bradyrhizobium	0.057142857
T5	Xanthobacter	0.057142857
T2	Xanthobacter	0.114285714
T1	Sorangium	0.142857143
T4	Sorangium	0.2
T3	Xanthobacter	0.30952381
T4	Russula	0.342857143
T5	Bradyrhizobium	0.342857143
T5	Burkholderia	0.342857143
T5	Sorangium	0.342857143
T6	Burkholderia	0.342857143
T6	Russula	0.342857143
T6	Sorangium	0.342857143
T2	Russula	0.4
T3	Bradyrhizobium	0.484848485
T3	Burkholderia	0.484848485
T4	Xanthobacter	0.485714286
T3	Russula	0.588744589
T2	Burkholderia	0.628571429
T2	Sorangium	0.628571429
T5	Russula	0.685714286
T3	Sorangium	0.699134199
T1	Russula	0.785714286
T6	Xanthobacter	0.885714286
T1	Bradyrhizobium	1
T1	Xanthobacter	1
T6	Bradyrhizobium	1

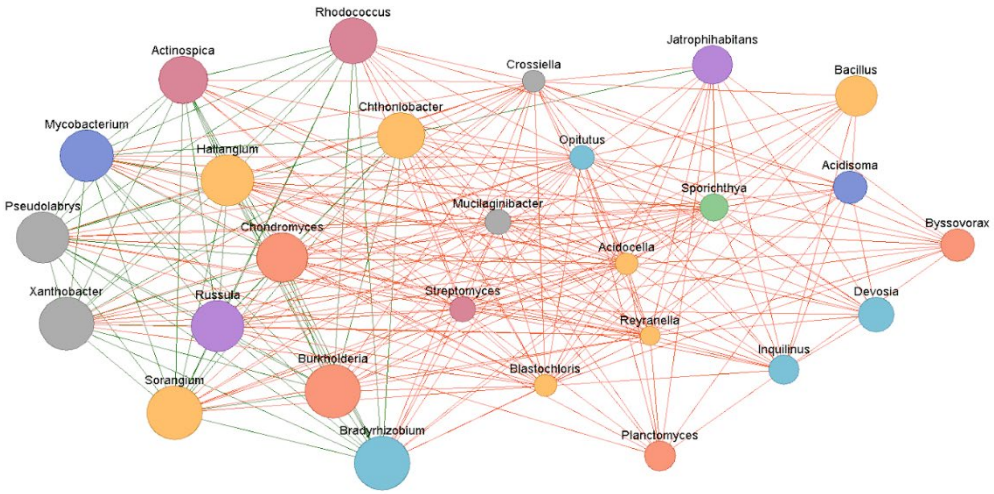

Figure 3.8: Co-occurrence network of taxonomic groups at the genus rank. Taxa shown for groups with a minimum threshold of 1% reads assigned, prevalence range of 10-100% of samples for which class rank must be considered present, and minimum probability of 70% with which two classes co-occur in samples. Edges depict co-occurrence in green and anti-occurrence and node size relative to abundance.

3.5 Discussion

Soil metatranscriptomes reflected a strong shift in microbial community composition with the warming treatment (Figure 3.3). The community structure, however, depicted no significant differences across the time series. This result was opposed to our expectations. The expectations derived from changes in nutrient availability (e.g., carbon), soil moisture and pH with the seasonal variation. Although the community structure was similar across the time series, it is possible that changes occurred at the functional level. In fact, the two previous chapters highlight that most changes occurred at the functional level. Another possibility involved the sequencing efforts, which sequencing depth resulted in a wide range for the metatranscriptomes. Also, the unequal number of replicates limited our statistical resolution to identify significant differences in low abundant taxa within the time series. As result, we focused our findings around the core taxonomic groups.

Changes associated to the warming effect were linked to soil-dwelling organisms playing an important role as symbionts, saprotrophs, or antibiotic producers. These organisms included both bacteria and fungi. Most bacterial changes were linked to Proteobacteria, Acidobacteria, and Actinobacteria. In turn, the fungal shift was linked to Basidiomycetes. At a lower resolution, the core taxonomic groups included Bradyrhizobium, Sorangium, Burkholderia, Xanthobacterer, and Russula (Figure 3.5). Members associated to Bradyrhizobium has been described as nitrogen-fixing symbionts, which enhanced plant growth with production of several molecules (Hayat et al., 2010). These molecules include auxins, cytokinins, abscisic acids, lumichrome, rhiboflavin, lipochitooligosaccharides and vitamins. The abundance of Bradyrhizobium increased with respect to the warming treatment at the Prospect Hill site.

Burkholderia is another symbiont and figured among most abundant taxa. Previous studies described this group as nitrogen fixers (Estrada-De Los Santos et al., 2001; Vandamme et al., 2002). At Prospect Hill site, the overall abundance of Burkholderia decrease with respect to the temperature effect. The seasonal effect did not reflect a clear pattern, however, an increase in abundance within the heated plots was recorded in the last two time points. A physiological characteristic of Burkholderia members is the flagella (Hofer, 2018). As described above, the precipitation and soil moisture in the last two time points is higher. Then, it is possible that bacterial members associated to Burkholderia increase as result of elevated soil moisture which favor the motility of the soil bacterium.

Russula was the only fungal group detected among the most abundant taxa. This group has been described as mycorrhizal fungi with high preference to birch forest (Geml et al., 2010; Girlanda et al., 2006). Based on the warming effect, the fungi group depicted a decrease in abundance. A similar trend has been recorded in other studies at the Harvard Forest (DeAngelis et al., 2015; Melillo et al., 2017). Changes in the time series were only detected in the controls.

As expected, the abundance of Russula peaked during the warmest time points and gradually decrease in the subsequent time points.

Warming resulted in the overrepresentation of saprophytic bacteria. As detected in the soil metatranscriptomes, the bacterial group Sorangium increase in abundance with warming. This deltaproteobacteria member is characterized to produce spores under unfavorable conditions (Mohr et al., 2018). In other studies, members isolated from soil has been described as saprotrophs and producer of antifungal metabolites (Pradella et al., 2002; Schneiker et al., 2007). Similarly, the Xanthobacter group reflected significant shifts with the warming treatment. Besides saprophytic bacteria, the Xanthobacter group is known to harbor nitrogen fixing members (Line, 1997). Abundance of both bacteria group showed a variable pattern across the time series (Figure 3.7).

A co-occurrence network analysis reflected a much more complex dynamic among taxonomic groups at the Prospect Hill site (Figure 3.8). Interestingly, the most abundant soil microbes depicted a positive correlation among them, whereas the low abundant taxa reflected a negative correlation. The co-occurrence network seems capture some common and expected biological relations. For instance, soil microbes previously described as bacterial symbionts correlated positively among them. However, the Russula group reflected an anti-occurrence with bacterial groups with antimicrobial properties. The negative correlation among low abundance soil-dwelling groups may arise as result of nutrient limitation. In fact, the long-term warming treatment has resulted in limitation of soil organic carbon (Melillo et al., 2011).

Taxonomic analysis based on rRNA represent a major challenge. We successfully annotated the soil metatranscriptomes with the tradeoff that only one percent of the overall sequences were annotated. The RiboTagger algorithm proved to be robust to annotate soil-dwelling microorganisms based on the V4 region. This included both eukaryotes (18S) and

bacteria (16S). Although the rRNA annotation was challenging, the results matched with previous metatranscriptomics analyses based on messenger RNA (mRNA). However, we would recommend that for metatranscriptomic analysis, especially for soil, to target the mRNA fraction. If possible, we highly recommend removing the rRNA. In previous experiments, this approach better captured microbial biodiversity of taxonomic groups across domain (i.e., viruses, archaea, and viruses). Furthermore, the mRNA profiles can provide insights of the functional activity of microbial communities. Also, the number of bioinformatic tools to mine the mRNA data is greater and continuously optimized, but otherwise with the rRNA data. When taken the above all together, metatranscriptomics must target the mRNA for a more robust approach in capturing activity in natural communities.

BIBLIOGRAPHY

Allison, S.D., Wallenstein, M.D., and Bradford, M.A. (2010). Soil-carbon response to warming dependent on microbial physiology. Nat. Geosci. *3*, 336–340.

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data.

Arndt, D.S., Blunden, J., and Hartfield, G. (2018). STATE OF THE CLIMATE IN 2017. Bull. Am. Meteorol. Soc. *99*, S1+.

Bailly, J., Fraissinet-Tachet, L., Verner, M.-C., Debaud, J.-C., Lemaire, M., Wésolowski-Louvel, M., and Marmeisse, R. (2007). Soil eukaryotic functional diversity, a metatranscriptomic approach. ISME J. *1*, 632–642.

Bardgett, R.D., Freeman, C., and Ostle, N.J. (2008). Microbial contributions to climate change through carbon cycle feedbacks. ISME J. *2*, 805–814.

Bastian, M., and Heymann, S. (2009). Gephi : An Open Source Software for Exploring and Manipulating Networks. ICWSM.

Benbi, D.K., Boparai, A.K., and Brar, K. (2014). Decomposition of particulate organic matter is more sensitive to temperature than the mineral associated organic matter. Soil Biol. Biochem. *70*, 183–192.

Berben, T., Overmars, L., Sorokin, D.Y., and Muyzer, G. (2019). Diversity and Distribution of Sulfur Oxidation-Related Genes in Thioalkalivibrio, a Genus of Chemolithoautotrophic and Haloalkaliphilic Sulfur-Oxidizing Bacteria. Front. Microbiol. *10*.

Bergmann, G.T., Bates, S.T., Eilers, K.G., Lauber, C.L., Caporaso, J.G., Walters, W.A., Knight, R., and Fierer, N. (2011). The under-recognized dominance of Verrucomicrobia in soil bacterial communities. Soil Biol. Biochem. *43*, 1450–1455.

Bertaux, J., Schmid, M., Hutzler, P., Hartmann, A., Garbaye, J., and Frey-Klett, P. (2005). Occurrence and distribution of endobacteria in the plant-associated mycelium of the ectomycorrhizal fungus Laccaria bicolor S238N. Environ. Microbiol. *7*, 1786–1795.

Bonfante, P., and Desirò, A. (2017). Who lives in a fungus? The diversity, origins and functions of fungal endobacteria living in Mucoromycota. ISME J. *11*, 1727–1735.

Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and Sensitive Protein Alignment using DIAMOND. Nat. Methods *12*.

Bushnell, B. (2016). BBtools.

Carney, K.M., Hungate, B.A., Drake, B.G., and Megonigal, J.P. (2007). Altered soil microbial community at elevated CO2 leads to loss of soil carbon. Proc. Natl. Acad. Sci. *104*, 4990–4995.

Cederlund, H., Wessén, E., Enwall, K., Jones, C.M., Juhanson, J., Pell, M., Philippot, L., and Hallin, S. (2014). Soil carbon quality and nitrogen fertilization structure bacterial communities with predictable responses of major bacterial phyla. Appl. Soil Ecol. *84*, 62–68.

Chang, Y., Land, M., Hauser, L., Chertkov, O., Rio, T.G.D., Nolan, M., Copeland, A., Tice, H., Cheng, J.-F., Lucas, S., et al. (2011). Non-contiguous finished genome sequence and contextual data of the filamentous soil bacterium Ktedonobacter racemifer type strain (SOSP1-21T). Stand. Genomic Sci. *5*, 97.

Chen, Y., Mccarthy, D., Robinson, M., and Smyth, G.K. (2015). edgeR : differential expression analysis of digital gene expression data User ' s Guide.

Choby, J.E., and Skaar, E.P. (2016). Heme Synthesis and Acquisition in Bacterial Pathogens. J. Mol. Biol. *428*, 3408–3428.

Classen, A.T., Sundqvist, M.K., Henning, J.A., Newman, G.S., Moore, J.A.M., Cregger, M.A., Moorhead, L.C., and Patterson, C.M. (2015). Direct and indirect effects of climate change on soil microbial and soil microbial-plant interactions: What lies ahead? Ecosphere *6*, art130.

Correa, E., Carvalhais, L., Utida, M., Oliveira, C., and Scotti, M. (2016). Effect of plant species on P cycle-related microorganisms associated with litter decomposition and P soil availability: implications for agroforestry management. IForest - Biogeosciences For. *9*, 294–302.

Crowther, T.W., Thomas, S.M., Maynard, D.S., Baldrian, P., Covey, K., Frey, S.D., van Diepen, L.T.A., and Bradford, M.A. (2015). Biotic interactions mediate soil microbial feedbacks to climate change. Proc. Natl. Acad. Sci. *112*, 7033–7038.

Cui, Y., Fang, L., Guo, X., Wang, X., Wang, Y., Li, P., Zhang, Y., and Zhang, X. (2018). Responses of soil microbial communities to nutrient limitation in the desert-grassland ecological transition zone. Sci. Total Environ. *642*, 45–55.

Curtin, D., Beare, M.H., and Hernandez-Ramirez, G. (2012). Temperature and Moisture Effects on Microbial Biomass and Soil Organic Matter Mineralization. Soil Sci. Soc. Am. J. *76*, 2055–2067.

DeAngelis, K.M., Pold, G., Topçuoğlu, B.D., van Diepen, L.T.A., Varney, R.M., Blanchard, J.L., Melillo, J., and Frey, S.D. (2015). Long-term forest soil warming alters microbial communities in temperate forest soils. Front. Microbiol. *6*.

Desirò, A., Hao, Z., Liber, J.A., Benucci, G.M.N., Lowry, D., Roberson, R., and Bonito, G. (2018). Mycoplasma -related endobacteria within Mortierellomycotina fungi: diversity, distribution and functional insights into their lifestyle. ISME J. *12*, 1743.

Dungait, J.A.J., Hopkins, D.W., Gregory, A.S., and Whitmore, A.P. (2012). Soil organic matter turnover is governed by accessibility not recalcitrance. Glob. Change Biol. *18*, 1781–1796.

Eichorst, S.A., Kuske, C.R., and Schmidt, T.M. (2011). Influence of Plant Polymers on the Distribution and Cultivation of Bacteria in the Phylum Acidobacteria. Appl. Environ. Microbiol. *77*, 586–596.

Estrada-De Los Santos, P., Bustillos-Cristales, R., and Caballero-Mellado, J. (2001). Burkholderia, a Genus Rich in Plant-Associated Nitrogen Fixers with Wide Environmental and Geographic Distribution. Appl. Environ. Microbiol. *67*, 2790–2798.

Faller, M., Matsunaga, M., Yin, S., Loo, J.A., and Guo, F. (2007). Heme is involved in microRNA processing. Nat. Struct. Mol. Biol. *14*, 23–29.

Fierer, N., and Jackson, R.B. (2006). The Diversity and Biogeography of Soil Bacterial Communities. Proc. Natl. Acad. Sci. U. S. A. *103*, 626–631.

Fierer, N., Bradford, M.A., and Jackson, R.B. (2007). Toward an ecological classification of soil bacteria. Ecology *88*, 1354–1364.

Fierer, N., Lauber, C.L., Ramirez, K.S., Zaneveld, J., Bradford, M.A., and Knight, R. (2012). Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. ISME J. *6*, 1007–1017.

Frey, S.D., Drijber, R., Smith, H., and Melillo, J. (2008). Microbial biomass, functional capacity, and community structure after 12 years of soil warming. Soil Biol. Biochem. *40*, 2904–2907.

Gans, J., Wolinsky, M., and Dunbar, J. (2005). Computational improvements reveal great bacterial diversity and high metal toxicity in soil. Science *309*, 1387–1390.

Geml, J., Laursen, G.A., Herriott, I.C., McFarland, J.M., Booth, M.G., Lennon, N., Nusbaum, H.C., and Taylor, D.L. (2010). Phylogenetic and ecological analyses of soil and sporocarp DNA sequences reveal high diversity and strong habitat partitioning in the boreal ectomycorrhizal genus Russula (Russulales; Basidiomycota). New Phytol. *187*, 494–507.

Ghosh, W., and Dam, B. (2009). Biochemistry and molecular biology of lithotrophic sulfur oxidation by taxonomically and ecologically diverse bacteria and archaea. FEMS Microbiol. Rev. *33*, 999–1043.

Gielnik, A., Pechaud, Y., Huguenot, D., Cébron, A., Riom, J.-M., Guibaud, G., Esposito, G., and van Hullebusch, E.D. (2019). Effect of digestate application on microbial respiration and bacterial communities' diversity during bioremediation of weathered petroleum hydrocarbons contaminated soils. Sci. Total Environ. *670*, 271–281.

Girlanda, M., Selosse, M.A., Cafasso, D., Brilli, F., Delfine, S., Fabbian, R., Ghignone, S., Pinelli, P., Segreto, R., Loreto, F., et al. (2006). Inefficient photosynthesis in the Mediterranean orchid Limodorum abortivum is mirrored by specific association to ectomycorrhizal Russulaceae. Mol. Ecol. *15*, 491–504.

Graff, A., and Stubner, S. (2003). Isolation and molecular characterization of thiosulfate-oxidizing bacteria from an Italian rice field soil. Syst. Appl. Microbiol. Stuttg. *26*, 445–452.

Gruwell, M.E., Morse, G.E., and Normark, B.B. (2007). Phylogenetic congruence of armored scale insects (Hemiptera: Diaspididae) and their primary endosymbionts from the phylum Bacteroidetes. Mol. Phylogenet. Evol. *44*, 267–280.

Hayatsu, M., Tago, K., and Saito, M. (2008). Various players in the nitrogen cycle: Diversity and functions of the microorganisms involved in nitrification and denitrification. Soil Sci. Plant Nutr. *54*, 33–45.

Heijden, M.G.A.V.D., Bardgett, R.D., and Straalen, N.M.V. (2008). The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. Ecol. Lett. *11*, 296–310.

Heimann, M., and Reichstein, M. (2008). Terrestrial ecosystem carbon dynamics and climate feedbacks. Nature *451*, 289–292.

Hibbett, D.S. (2006). A phylogenetic overview of the Agaricomycotina. Mycologia *98*, 917–925.

Hofer, U. (2018). Bacterial physiology: It's a wrap for *Burkholderia* flagella. Nat. Rev. Microbiol. *16*, 65.

Hu, Y., Zheng, Q., Zhang, S., Noll, L., and Wanek, W. (2018). Significant release and microbial utilization of amino sugars and d-amino acid enantiomers from microbial cell wall decomposition in soils. Soil Biol. Biochem. *123*, 115–125.

Huson, D.H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., and Tappu, R. (2016). MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. PLOS Comput. Biol. *12*, e1004957.

Janssen, P.H. (2006). Identifying the Dominant Soil Bacterial Taxa in Libraries of 16S rRNA and 16S rRNA Genes. Appl Env. Microbiol *72*, 1719–1728.

Jayasinghearachchi, H.S., and Seneviratne, G. (2004). A bradyrhizobial-Penicillium spp. biofilm with nitrogenase activity improves N2 fixing symbiosis of soybean. Biol. Fertil. Soils *40*, 432–434.

Jin, M., Guo, X., Zhang, R., Qu, W., Gao, B., and Zeng, R. (2019). Diversities and potential biogeochemical impacts of mangrove soil viruses. Microbiome *7*.

Jing, X., Sanders, N.J., Shi, Y., Chu, H., Classen, A.T., Zhao, K., Chen, L., Shi, Y., Jiang, Y., and He, J.-S. (2015). The links between ecosystem multifunctionality and above- and belowground biodiversity are mediated by climate. Nat. Commun. *6*, 8159.

Jordan, L.D., Zhou, Y., Smallwood, C.R., Lill, Y., Ritchie, K., Yip, W.T., Newton, S.M., and Klebba, P.E. (2013). Energy-dependent motion of TonB in the Gram-negative bacterial inner membrane. Proc. Natl. Acad. Sci. *110*, 11553–11558.

Jose, J., Snyder, J.E., and Kuhn, R.J. (2009). A structural and functional perspective of alphavirus replication and assembly. Future Microbiol. Lond. *4*, 837–856.

Kaspari, M., Garcia, M.N., Harms, K.E., Santana, M., Wright, S.J., and Yavitt, J.B. (2008). Multiple nutrients limit litterfall and decomposition in a tropical forest. Ecol. Lett. *11*, 35–43.

Kennedy, P., and Stajich, J. (2015). Twenty-first century mycology: a diverse, collaborative, and highly relevant science. New Phytol. *205*, 23–26.

King, C.E., and King, G.M. (2014). Description of Thermogemmatispora carboxidivorans sp. nov., a carbon-monoxide-oxidizing member of the class Ktedonobacteria isolated from a geothermally heated biofilm, and analysis of carbon monoxide oxidation by members of the class Ktedonobacteria. Int. J. Syst. Evol. Microbiol. *64*, 1244–1251.

Kleber, M., Nico, P.S., Plante, A., Filley, T., Kramer, M., Swanston, C., and Sollins, P. (2011). Old and stable soil organic matter is not necessarily chemically recalcitrant: implications for modeling concepts and temperature sensitivity. Glob. Change Biol. *17*, 1097–1107.

Kopylova, E., Noe, L., and Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics *28*, 3211–3217.

Koyama, A., Wallenstein, M.D., Simpson, R.T., and Moore, J.C. (2014). Soil bacterial community composition altered by increased nutrient availability in Arctic tundra soils. Front. Microbiol. *5*.

Kujawinski, E.B., and Behn, M.D. (2006). Automated Analysis of Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectra of Natural Organic Matter. Anal. Chem. *78*, 4363–4373.

Lau, E., Ahmad, A., Steudler, P.A., and Cavanaugh, C.M. (2007). Molecular characterization of methanotrophic communities in forest soils that consume atmospheric methane. FEMS Microbiol. Ecol. *60*, 490–500.

Lehembre, F., Doillon, D., David, E., Perrotto, S., Baude, J., Foulon, J., Harfouche, L., Vallon, L., Poulain, J., Silva, C.D., et al. (2013). Soil metatranscriptomics for mining eukaryotic heavy metal resistance genes. Environ. Microbiol. *15*, 2829–2840.

Line, M.A. (1997). A nitrogen-fixing consortia associated with the bacterial decay of a wooden pipeline. Lett. Appl. Microbiol. *25*, 220–224.

Lun, A.T.L., Chen, Y., and Smyth, G.K. (2016). It's DE-licious: A Recipe for Differential Expression Analyses of RNA-seq Experiments Using Quasi-Likelihood Methods in edgeR. In Statistical Genomics: Methods and Protocols, E. Mathé, and S. Davis, eds. (New York, NY: Springer New York), pp. 391–416.

Magoč, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics *27*, 2957–2963.

Mann, B.F., Chen, H., Herndon, E.M., Chu, R.K., Tolic, N., Portier, E.F., Chowdhury, T.R., Robinson, E.W., Callister, S.J., Wullschleger, S.D., et al. (2015). Indexing Permafrost Soil Organic Matter Degradation Using High-Resolution Mass Spectrometry. PLOS ONE *10*, e0130557.

Maron, P.-A., Sarr, A., Kaisermann, A., Lévêque, J., Mathieu, O., Guigue, J., Karimi, B., Bernard, L., Dequiedt, S., Terrat, S., et al. (2018). High Microbial Diversity Promotes Soil Ecosystem Functioning. Appl. Environ. Microbiol. *84*, e02738-17.

Melillo, J.M. (2002). Soil Warming and Carbon-Cycle Feedbacks to the Climate System. Science *80*, 2173–2176.

Melillo, J.M. (2011). Soil warming, carbon-nitrogen interactions, and forest carbon budgets.

Melillo, J.M., Butler, S., Johnson, J., Mohan, J., Steudler, P., Lux, H., Burrows, E., Bowles, F., Smith, R., Scott, L., et al. (2011). Soil warming, carbon–nitrogen interactions, and forest carbon budgets. Proc. Natl. Acad. Sci. U. S. A. *108*, 9508–9512.

Melillo, J.M., Frey, S.D., DeAngelis, K.M., Werner, W.J., Bernard, M.J., Bowles, F.P., Pold, G., Knorr, M.A., and Grandy, A.S. (2017). Long-term pattern and magnitude of soil carbon feedback to the climate system in a warming world. Science *358*, 101–105.

Miller, T.R., Delcher, A.L., Salzberg, S.L., Saunders, E., Detter, J.C., and Halden, R.U. (2010). Genome Sequence of the Dioxin-Mineralizing Bacterium Sphingomonas wittichii RW1. J. Bacteriol. *192*, 6101–6102.

Minor, E.C., Steinbring, C.J., Longnecker, K., and Kujawinski, E.B. (2012). Characterization of dissolved organic matter in Lake Superior and its watershed using ultrahigh resolution mass spectrometry. Org. Geochem. *43*, 1–11.

Mohr, K.I., Wolf, C., Nübel, U., Szafrańska, A.K., Steglich, M., Hennessen, F., Gemperlein, K., Kämpfer, P., Martin, K., Müller, R., et al. (2018). A polyphasic approach leads to seven new species of the cellulose-decomposing genus Sorangium, Sorangium ambruticinum sp. nov., Sorangium arenae sp. nov., Sorangium bulgaricum sp. nov., Sorangium dawidii sp. nov., Sorangium kenyense sp. nov., Sorangium orientale sp. nov. and Sorangium reichenbachii sp. nov. Int. J. Syst. Evol. Microbiol. *68*, 3576–3586.

Moir, J.W.B., and Wood, N.J. (2001). Nitrate and nitrite transport in bacteria: Cell. Mol. Life Sci. *58*, 215–224.

Neilands, J.B. (1995). Siderophores: Structure and Function of Microbial Iron Transport Compounds. J. Biol. Chem. *270*, 26723–26726.

Ohm, R.A., Feau, N., Henrissat, B., Schoch, C.L., Horwitz, B.A., Barry, K.W., Condon, B.J., Copeland, A.C., Dhillon, B., Glaser, F., et al. (2012). Diverse Lifestyles and Strategies of Plant Pathogenesis Encoded in the Genomes of Eighteen Dothideomycetes Fungi. PLOS Pathog. *8*, e1003037.

Okie, J.G., Van Horn, D.J., Storch, D., Barrett, J.E., Gooseff, M.N., Kopsova, L., and Takacs-Vesbach, C.D. (2015). Niche and metabolic principles explain patterns of diversity and distribution: theory and a case study with soil bacterial communities. Proc. R. Soc. B Biol. Sci. *282*, 20142630.

Okuda, S., Yamada, T., Hamajima, M., Itoh, M., Katayama, T., Bork, P., Goto, S., and Kanehisa, M. (2008). KEGG Atlas mapping for global analysis of metabolic pathways. Nucleic Acids Res. *36*, W423-426.

Partida-Martinez, L.P., and Hertweck, C. (2005). Pathogenic fungus harbours endosymbiotic bacteria for toxin production. Nature *437*, 884.

Pascual, J., Wüst, P.K., Geppert, A., Foesel, B.U., Huber, K.J., and Overmann, J. (2015). Novel isolates double the number of chemotrophic species and allow the first description of higher taxa in Acidobacteria subdivision 4. Syst. Appl. Microbiol. *38*, 534–544.

Peterjohn, W.T., Melillo, J.M., Bowles, F.P., and Steudler, P.A. (1993). Soil warming and trace gas fluxes: experimental design and preliminary flux results. Oecologia *93*, 18–24.

Peters, M.K., Hemp, A., Appelhans, T., Becker, J.N., Behler, C., Classen, A., Detsch, F., Ensslin, A., Ferger, S.W., Frederiksen, S.B., et al. (2019). Climate–land-use interactions shape tropical mountain biodiversity and ecosystem functions. Nature *568*, 88.

Pianka, E.R. (1970). On r- and K-Selection. Am. Nat. *104*, 592–597.

Pisani, O., Frey, S.D., Simpson, A.J., and Simpson, M.J. (2015). Soil warming and nitrogen deposition alter soil organic matter composition at the molecular-level. Biogeochemistry *123*, 391–409.

Pold, G., Billings, A.F., Blanchard, J.L., Burkhardt, D.B., Frey, S.D., Melillo, J.M., Schnabel, J., Diepen, L.T.A. van, and DeAngelis, K.M. (2016). Long-Term Warming Alters Carbohydrate Degradation Potential in Temperate Forest Soils. Appl Env. Microbiol *82*, 6518–6530.

Pold, G., Grandy, A.S., Melillo, J.M., and DeAngelis, K.M. (2017). Changes in substrate availability drive carbon cycle response to chronic warming. Soil Biol. Biochem. *110*, 68–78.

Pradella, S., Hans, A., Spröer, C., Reichenbach, H., Gerth, K., and Beyer, S. (2002). Characterisation, genome size and genetic manipulation of the myxobacterium Sorangium cellulosum So ce56. Arch. Microbiol. *178*, 484–492.

Pries, C.E.H., Castanha, C., Porras, R.C., and Torn, M.S. (2017). The whole-soil carbon flux in response to warming. Science *355*, 1420–1423.

R Core Team, and R Development Core Team (2008). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).

Rich, P.R. (2003). The molecular machinery of Keilin's respiratory chain. Biochem. Soc. Trans. *31*, 1095–1105.

RStudio Team (2016). RStudio: Integrated Development Environment for R (Boston, MA: RStudio, Inc.).

Ruess, L., Michelsen, A., Schmidt, I.K., and Jonasson, S. (1999). Simulated climate change affecting microorganisms, nematode density and biodiversity in subarctic soils. Plant Soil *212*, 63–73.

Schmidt, M.W.I., Torn, M.S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I.A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D.A.C., et al. (2011). Persistence of soil organic matter as an ecosystem property. Nature *478*, 49–56.

Schneiker, S., Perlova, O., Kaiser, O., Gerth, K., Alici, A., Altmeyer, M.O., Bartels, D., Bekel, T., Beyer, S., Bode, E., et al. (2007). Complete genome sequence of the myxobacterium Sorangium cellulosum. Nat. Biotechnol. *25*, 1281–1289.

Schroter, D., Wolters, V., and De Ruiter, P.C. (2003). C and N mineralisation in the decomposer food webs of a European forest transect. Oikos *102*, 294–308.

Schulz, F., Alteio, L., Goudeau, D., Ryan, E.M., Yu, F.B., Malmstrom, R.R., Blanchard, J., and Woyke, T. (2018). Hidden diversity of soil giant viruses. Nat. Commun. *9*, 4881.

Sharma, M., Schmid, M., Rothballer, M., Hause, G., Zuccaro, A., Imani, J., Kämpfer, P., Domann, E., Schäfer, P., Hartmann, A., et al. (2008). Detection and identification of bacteria intimately associated with fungi of the order Sebacinales. Cell. Microbiol. *10*, 2235–2246.

Simon, C., Roth, V.-N., Dittmar, T., and Gleixner, G. (2018). Molecular Signals of Heterogeneous Terrestrial Environments Identified in Dissolved Organic Matter: A Comparative Analysis of Orbitrap and Ion Cyclotron Resonance Mass Spectrometers. Front. Earth Sci. *6*.

Sjursen, H., Michelsen, A., and Jonasson, S. (2005). Effects of long-term soil warming and fertilisation on microarthropod abundances in three sub-arctic ecosystems. Appl. Soil Ecol. *30*, 148–161.

Smit, E., Leeflang, P., Gommans, S., Broek, J. van den, Mil, S. van, and Wernars, K. (2001). Diversity and Seasonal Fluctuations of the Dominant Members of the Bacterial Soil Community in a Wheat Field as Determined by Cultivation and Molecular Methods. Appl Env. Microbiol *67*, 2284–2291.

Spatafora, J.W., Owensby, C.A., Douhan, G.W., Boehm, E.W.A., and Schoch, C.L. (2012). Phylogenetic placement of the ectomycorrhizal genus Cenococcum in Gloniaceae (Dothideomycetes). Mycologia *104*, 758–765.

Staley, J.T., and Konopka, A. (1985). Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. Annu. Rev. Microbiol. *39*, 321–346.

Summerbell, R.C. (2005). Root endophyte and mycorrhizosphere fungi of black spruce, Picea mariana, in a boreal forest habitat: influence of site factors on fungal distributions. Stud. Mycol. *53*, 121–145.

Taggart, M., Heitman, J.L., Shi, W., and Vepraskas, M. (2012). Temperature and Water Content Effects on Carbon Mineralization for Sapric Soil Material. Wetlands *32*, 939–944.

Tange, O. (2011). GNU Parallel: the command-line power tool. Login USENIX Mag. *36*, 42–47.

Tfaily, M.M., Chu, R.K., Tolic, N., Roscioli, K.M., Anderton, C.R., Pasa-Tolic, L., Robinson, E.W., and Hess, N.J. (2015). Advanced Solvent Based Methods for Molecular Characterization of Soil Organic Matter by High-Resolution Mass Spectrometry. Anal. Chem. *87*, 5206–5215.

Tfaily, M.M., Chu, R.K., Toyoda, J., Tolic, N., Robinson, E.W., Pasa-Tolic, L., and Hess, N.J. (2017a). Sequential extraction protocol for organic matter from soils and sediments using high resolution mass spectrometry. Anal. Chim. Acta *972*, 54–61.

Tfaily, M.M., Chu, R.K., Toyoda, J., Tolic, N., Robinson, E.W., Pasa-Tolic, L., and Hess, N.J. (2017b). Sequential extraction protocol for organic matter from soils and sediments using high resolution mass spectrometry. Anal. Chim. Acta *972*, 54–61.

Tong, J., Miaowen, C., Juhui, J., Jinxian, L., and Baofeng, C. (2017). Endophytic fungi and soil microbial community characteristics over different years of phytoremediation in a copper tailings dam of Shanxi, China. Sci. Total Environ. *574*, 881–888.

Trumbore, S.E. (1997). Potential responses of soil organic carbon to global environmental change. Proc. Natl. Acad. Sci. *94*, 8284–8291.

Upton, R.N., Bach, E.M., and Hofmockel, K.S. (2019). Spatio-temporal microbial community dynamics within soil aggregates. Soil Biol. Biochem. *132*, 58–68.

Uroz, S., Ioannidis, P., Lengelle, J., Cébron, A., Morin, E., Buée, M., and Martin, F. (2013). Functional Assays and Metagenomic Analyses Reveals Differences between the Microbial Communities Inhabiting the Soil Horizons of a Norway Spruce Plantation. PLOS ONE *8*, e55929.

Vandamme, P., Goris, J., Chen, W.-M., de Vos, P., and Willems, A. (2002). Burkholderia tuberum sp. nov. and Burkholderia phymatum sp. nov., Nodulate the Roots of Tropical Legumes. Syst. Appl. Microbiol. *25*, 507–512.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D., Paulsen, I., Nelson, K.E., Nelson, W., et al. (2004). Environmental genome shotgun sequencing of the Sargasso Sea. Science *304*, 66–74.

Viswanath, G.K., and Pillai, S.C. (1972). Influence of moisture on soil aggregation. Curr. Sci. *41*, 547–553.

Wang, Z., Binder, M., Schoch, C.L., Johnston, P.R., Spatafora, J.W., and Hibbett, D.S. (2006). Evolution of helotialean fungi (Leotiomycetes, Pezizomycotina): A nuclear rDNA phylogeny. Mol. Phylogenet. Evol. *41*, 295–312.

Ward, C.P., Nalven, S.G., Crump, B.C., Kling, G.W., and Cory, R.M. (2017). Photochemical alteration of organic carbon draining permafrost soils shifts microbial metabolic pathways and stimulates respiration. Nat. Commun. *8*, 772.

Wei, H., Xu, Q., Taylor, L.E., Baker, J.O., Tucker, M.P., and Ding, S.-Y. (2009). Natural paradigms of plant cell wall degradation. Curr. Opin. Biotechnol. *20*, 330–338.

Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis (New York: Springer-Verlag).

Wieder, W.R., Bonan, G.B., and Allison, S.D. (2013). Global soil carbon projections are improved by modelling microbial processes. Nat. Clim. Change *3*, 909–912.

Wiesmeier, M., Urbanski, L., Hobley, E., Lang, B., von Lützow, M., Marin-Spiotta, E., van Wesemael, B., Rabot, E., Ließ, M., Garcia-Franco, N., et al. (2019). Soil organic carbon storage as a key function of soils - A review of drivers and indicators at various scales. Geoderma *333*, 149–162.

Wild, B., Schnecker, J., Alves, R.J.E., Barsukov, P., Bárta, J., Čapek, P., Gentsch, N., Gittel, A., Guggenberger, G., Lashchinskiy, N., et al. (2014). Input of easily available organic C and N stimulates microbial decomposition of soil organic matter in arctic permafrost soil. Soil Biol. Biochem. *75*, 143–151.

Wohl, D.L., Arora, S., and Gladstone, J.R. (2004). FUNCTIONAL REDUNDANCY SUPPORTS BIODIVERSITY AND ECOSYSTEM FUNCTION IN A CLOSED AND CONSTANT ENVIRONMENT. Ecology *85*, 1534–1540.

Wolinsky, E., and Rynearson, T.K. (2015). Mycobacteria in Soil and Their Relation to Disease-Associated Strains1, 2, 3. Am. Rev. Respir. Dis.

Wu, X., Wu, L., Liu, Y., Zhang, P., Li, Q., Zhou, J., Hess, N.J., Hazen, T.C., Yang, W., and Chakraborty, R. (2018). Microbial Interactions With Dissolved Organic Matter Drive Carbon Dynamics and Community Succession. Front. Microbiol. *9*.

Xie, C., Goi, C.L.W., Huson, D.H., Little, P.F.R., and Williams, R.B.H. (2016). RiboTagger: fast and unbiased 16S/18S profiling using whole community shotgun metagenomic or metatranscriptome surveys. BMC Bioinformatics *17*.

Yang, P., Zhang, M., and Elsas, J.D. van (2017). Role of flagella and type four pili in the co-migration of Burkholderia terrae BS001 with fungal hyphae through soil. Sci. Rep. *7*, 2997.

Yin, H., Li, Y., Xiao, J., Xu, Z., Cheng, X., and Liu, Q. (2013). Enhanced root exudation stimulates soil nitrogen transformations in a subalpine coniferous forest under experimental warming. Glob. Change Biol. *19*, 2158–2167.

Zhou, J., He, Z., Yang, Y., Deng, Y., Tringe, S.G., and Alvarez-Cohen, L. (2015). High-Throughput Metagenomic Technologies for Complex Microbial Community Analysis: Open and Closed Formats. MBio *6*, e02288-14.

Zhou, X., Lindsay, H., and Robinson, M.D. (2013). Robustly detecting differential expression in RNA sequencing data using observation weights. ArXiv13123382 Q-Bio Stat.