DE-SC0010570

Mapping soil carbon from cradle to grave: drafting a molecular blueprint for C transformation from roots to stabilized soil organic C

Final Technical Report

Period Covered: September 2013-August 2017

January 2018

Principal Investigator: Mary K. Firestone
Department of Environmental Science, Policy, and Management
University of California, Berkeley

Prepared for the U.S. Department of Energy Office of Science under Grant DE-SC0010570

Project Title: Mapping soil carbon from cradle to grave: drafting a molecular blueprint for C transformation from roots to stabilized soil organic C

Final Technical Report: Grant no. DE-SC0010570

Summary

The soil surrounding plant roots, the rhizosphere, has long been recognized as a zone of great functional importance to plants and the terrestrial ecosystems they inhabit. The primary objective of this research project was to determine how organic carbon (C) decomposition and stabilization processes in soil are impacted by the interactions between plant roots and the soil microbial community. The project addressed three hypotheses:

H1: The microbiomes of the rhizosphere and detritosphere undergo a functional succession driven by the molecular composition and quantity of root-derived C.

H2: Elevated CO_2 impacts the function and succession of rhizosphere communities thus altering the fate of root-derived C.

H3: Microbial metabolism of root derived C is a critical controller of the accumulation of organic C in the mineral-associated soil pool.

Researchers combined stable isotope approaches with metagenomic analyses in order to map the flow of C from roots to specific organisms within the rhizosphere. These analyses allowed us to assess the metabolic capabilities and functional profiles of the organisms using root carbon.

Outcomes

The interconnected rhizosphere: using network analysis to identify potential microbe-microbe interactions throughout the lifespan of a common annual grass

We examined the co-occurrence/co-exclusion patterns of bacteria in the rhizosphere soil of *Avena fatua,* a common Mediterranean annual grass, and compared them to patterns in the surrounding soil. Random matrix theory (RMT) was employed for network analysis of bacterial taxa (16S rRNA sequencing) over two growing seasons. Our results revealed that bacterial networks in rhizosphere soil were substantially more complex than networks in the surrounding soils. We showed that the development of these networks tracks bacterial succession in rhizosphere, where the complexity of the bacterial networks increased as the plants grew. The surrounding soils remained relatively static over the first season of plant growth, and slightly declined in complexity over the second season of plant growth. Within the rhizosphere networks, groups of highly connected "modules" formed over time. The high inter-connectivity in the rhizosphere and lack of connectivity in the surrounding soil suggest fundamental properties about each soil habitat. Roots provide bacterial community niches in which assemblages have an enhanced potential to interact; while the surrounding soil is composed of more fragmented or isolated niches that support an array of small, disparate assemblages with low interaction potential (Shi et al., 2016).

Rhizosphere metatranscriptomics

To draft the blueprint of plant C (exudates and root debris) cycling in the rhizosphere, we sequenced community transcriptomes of both rhizosphere and bulk soil communities sampled from microcosms grown with *Avena fatua* seedlings, both with and without *Avena* root litter. These samples were collected at four different time points over a month of peak root growth in

replicate microcosms, and transcripts were then sequenced to determine how root exudates alter decomposition of both simple and complex C substrates. The resulting libraries are high quality, with approximately 90% of the total reads having quality scores >Q30, and containing only 20% rRNA reads on average. To determine if roots alter the expression of carbohydrate-active enzymes relevant to litter decomposition, we performed a read-based annotation using the Carbohydrate Active enZYme (CAZy) database. Initial results suggest that some CAZymes relevant to litter decomposition (e.g. laccases) are elevated in the presence of a root. In particular, the enzyme families AA1 (e.g., laccase) and AA7 (e.g., glucooligosaccharide and chitooligosaccharide oxidases) were elevated over time in the rhizosphere soil relative to bulk soil (>2x rlog fold). (Nuccio, in prep)

Single Cell Genomics

We conducted an experiment to study the changes of rhizosphere and endophyte communities (bacterial 16S and fungal ITS) over the course of 3 months as roots die. Single cells were isolated from the rhizosphere and endosphere of living roots (T0), as well as detrital roots (T4, 3 months after senescence). We sorted 189 cells from these samples, which represent 86 phylotypes (99% identity). After removing taxa that have a close relative with an already-sequenced genome, we identified 51 cells for single amplified genome sequencing. Three of these cells are potentially novel at the class or phylum level, and almost all the remaining cells are <97% similar to existing genomes. The single-cell genomes, plus two recently re-sequenced metagenomes, and 50 Hopland soil isolates will be used as references to map the transcriptomes, as well as 'fill in gaps' in our community metagenome. These genomes sequences will allow us to better understand taxa and functions that are enriched in the rhizosphere, especially exudate-users involved in complex plant polymer degradation.

Chip-SIP: Automated Probe Design

We have developed an automated probe design pipeline to accelerate the design of custom Chip-SIP arrays from next-generation sequencing dataset; in part to improve our own use of the method, but also to facilitate the application of Chip-SIP by the greater scientific community. This pipeline includes a probe specificity check to reduce cross-hybridization between probes. We've used this pipeline to create a master rhizosphere microarray that targets >3,000 OTUs (>100,000 probes) from the *Avena* rhizosphere and surrounding soil, from which Chip SIP probes can be selected for new samples.

Profiling of *Avena* Root Exudates

A sterile plant growth chamber (SPC) was developed and used to quantify Avena barbata exudates for weeks under elevated and ambient CO2 conditions. GC/MS methods were used to characterize the root exudates. Out of 125 different metabolites detected and identified, only seven metabolites were found to be significantly different between elevated and ambient CO_2 treatments according to a two way ANOVA and Tukey's HSD with a $p < 0.05$. All metabolites found, except for 3-hydroxy-3-methylglutaric acid, only differed significantly between the two CO_2 treatments at one time point. 3-hydroxy-3-methylglutaric acid differed significantly between the two CO_2 treatments in both weeks six and nine. (Estera, thesis 2017)

Composition of and Microbial Preference for Root Exudates Drive Rhizosphere Community Composition. Plants may regulate the composition of their rhizosphere to promote the growth of microorganisms that improve plant fitness in a given ecosystem. Using a combination of comparative genomics and exometabolomics, we show that pre-programmed developmental

processes in plants result in consistent patterns in the chemical composition of root exudates. This chemical succession in the rhizosphere interacts with microbial metabolite substrate preferences that are predictable from genome sequences. The combination of these plant exudation traits and microbial substrate uptake traits interact to yield the consistent successional patterns observed in the rhizosphere microbiome of an annual grass. This discovery provides a mechanistic underpinning for the process of rhizosphere microbial community assembly and provides an attractive direction for the manipulation of the rhizosphere microbiome for beneficial outcomes. (Zhalnina et al. in review).

Developing Proteomic Methods for Extraction and Mapping of Soil Carbon from Plants and Microbes

We have n effectively developed a method for protein extraction from Hopland soils and along with a well-defined soil metagenome and genome sequences from isolates have begun to effectively characterize the microbial proteomes from soil.

Since the use of stable isotope probes is one method to determine the functionally active portion of a microbial community, we amended Hopland soils with 13C labeled glucose to determine a) the overall proteomic profile of the microbes in the soil, and b) track the label into those proteins that are synthesized in response to the glucose. For the initial goal, we found a suite of proteins that had functions including Oxidative stress, Carbon metabolism, Sulfur metabolism, TCA Cycle, Trehalose metabolism, Protein Translation and the transporters of Lactate/Malate, Citrate, Chitin, Glucose, Xylose, Trehalose. These initial results point to an increase in glucose-related central metabolism, transport, and sulfur metabolism in mineral associated microbes.

Abundance of isolates in dry/wet season metagenomes of Hopland soil

To evaluate the metabolic potential and the representation of our isolated strains within the metagenome from Hopland soil we mapped a subset (181457200 reads in MG1 and 178288802 reads in MG2) of paired reads in MG1 (wet season, 3949672770 reads total) and MG2 (dry season, 6039960896 reads total) to 38 isolate genomes. We analyzed representation of the metabolic functions in the metagenomic reads mapped at different levels of identity to the isolate sequences. Our mapping with whole set of MG1 and MG2 reads will reveal the metabolic functions and genomic diversity of the community more thoroughly. Using the reads mapped with 0-100% of identity, we are studying the degree of strain-level variation of soil bacteria.

We have also examined key physiological/metabolic traits of the isolated bacteria, including generation times, exoenzyme production and substrate utilization capabilities. We have confirmed predicted minimum generation times estimated using codon usage bias of sequenced isolates through comparison with empirical observations and demonstrated accuracy of this approach to predict growth rates of uncultured taxa *in silico*.

Exoenzyme production

We have examined the distribution of genomic potential for exoenzyme production (glycoside hydrolases and polyphenol oxidases) in the sequenced genomes and experimentally confirmed their activities using nanostructure-initiator mass spectrometry (NIMS) enzyme and colorimetric assays. Secretomes of the selected isolates were then used to determine identities of the exoenzymes produced and evaluate the conditions under which these enzymes are induced or repressed.

Substrate utilization

The phylogenetic distributions of sugar and organic acid transporters were evaluated across our 38 sequenced isolates. We have confirmed preferences of the selected isolates to utilize specific substrates and the release of a variety of metabolites using untargeted exometabolomics following cultivation on a rich medium. To determine consumption of various components of *Avena* root exudates by our bacterial isolates and what metabolites are produced following consumption of exudates we have used liquid chromatography-mass spectrometry to compare metabolic profiles of media with Avena exudates and the same media post isolate growth.

Genome-scale metabolic models

To predict *in silico* metabolic traits of the microbial community in the studied soil we have reconstructed a draft of genome-scale models of four key microorganisms from the Hopland grassland using KBase. Phenotypic microarrays and exometabolomics were used to define substrate uptake/release of the modeled bacteria and gapfill metabolic models. Further uptake kinetics of substrates are being determined to curate fluxes and validate metabolic models. These models once reconciled with physiological data will be used to simulate plant-microbial interactions though metabolic fluxes in the microbial cell in response to plant exudation; and to inform trait-based model with key metabolic traits of the rhizosphere bacteria.

Metagenomic Analysis of Mediterranean grassland soil: assembly, binning and annotation

Metagenomic short read data from two soil samples collected after the dry summer season and at the peak of the wet season were analyzed using an assembly driven approach. Current sequencing from these two samples (MG1 &2) yielded a total of 0.77 Tbp of high quality sequence. This data was co-assembled using a newly developed iterative exhaustive assembly pipeline that scales to Tbp of short read data**.** The resulting assembly was binned using cross-sample coverage and tetranucleotide frequencies. The assembly was in 138,990 scaffolds, represented ~10% of the reads, had a total length of 0.3 Gbp and N50 scaffold length of 1.2 kbp. This assembly pipeline is currently being ported to KBase in collaboration with the KBase team. Taxonomic profiling of the assembly demonstrated that soil was dominated by Bacteria with several prominent phyla, with Archaea (Cren/Thaumarchaea) and Eukarya (Metazoa, Fungi) were also represented. Currently, genome/organism binning identified 66 bins (~50% of all scaffolds were binned). Bin completeness and purity were verified against ribosomal protein and single copy genes. 36 bins were identified as belonging to Actinobacteria and 20 bins to Proteobacteria the dominant phyla in this soil. Actinobacterial bins were primarily from Actinomycetales and Solirubrobacterales while Proteobacterial bins were primarily from Alpha-proteobacteria (Rhizobialess) that correspond closely to isolated representatives.

Carbon flow from roots to mineral surfaces

We completed a 3.5 month greenhouse experiment with 600 mineral "coupons", pure minerals isolated in 18 μm mesh and incubated in soil with growing *Avena barbata* plants in a 99 atom% $^{13}CO_2$ atmosphere. From Elemental Analysis/Isotope Ratio Mass Spectrometry (EA/IRMS), we see that pure minerals accumulate C over the growing season of *A. barbata*, with an increasing ratio of plant-derived ^{13}C (**Fig**). The amount of C and relative source of that C are dependent on mineralogy, with the most reactive mineral type, the iron (hydr)oxide ferrihydrite, trends toward accumulating the most C by mass and the non-reactive quartz accumulating the least. When normalized to surface area, however, quartz had the highest total C. We believe that, contrary to our initial hypothesis, minerals did not become fully "saturated" with C in the first growing season. This idea is further supported by an interesting finding when comparing pure minerals to natural minerals from the soil, which are already associated with C: in our control, no plant treatment, natural soil minerals accumulated less C than in the planted treatment. However, pure

minerals (ferrihydrite, quartz, and kaolinite) accumulated more C in the no plant control treatment, suggesting that the form of C and/or the processes controlling C association on a fresh mineral surface are different from those on a mineral surface that already has C associated with it. Through DNA extraction and quantification, we find that a small but potentially important fraction of C (15-150 ng DNA g^{-1} dry mineral) is microbial.

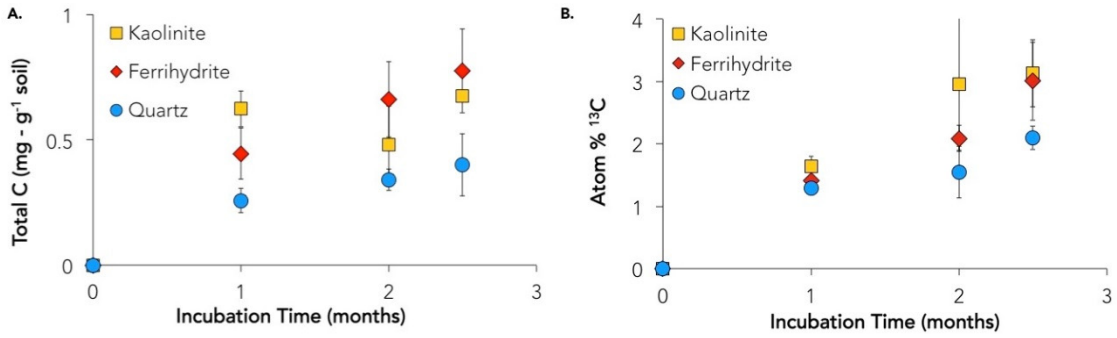

Figure. Total C (A) and atom % ^{13}C (B) accumulation on surfaces of different mineral types: kaolinite, ferrihydrite, and quartz. Atom % ^{13}C indicates plant-derived C, although the composition of that C is likely transformed by microbial processes.

A functional gene-oriented metagenomic analysis pipeline (FunGen-MAP) We developed a functional gene-oriented metagenomic analysis pipeline (FunGen-MAP) for rapid processing and annotation of shotgun metagenome sequencing reads. Briefly, the sequence reads are compared to reference functional gene models and sequences known to be associated with important geochemical and ecological processes. The developed FunGen-MAP consists of two major workflows: reference database preparation and sequencing read annotation. The validation of the FunGen-MAP pipeline is being conducted based on both the artificial metagenomes and project sequencing data. Based on the results we currently have, the pipeline shows high accuracy in terms of true and false positive ratios of annotations, and optimal performance on the read length range from 150 to 400bp. Meanwhile, the pipeline is considered to be robust in terms of its error tolerance performance, with the results showing that its accuracy was only slightly affected by the read error ratio up to 5%.

Publications:

- Nuccio, E. 2013. Molecular mechanisms involved in root-microbial interactions: effects on decomposition, Ph.D. Thesis, UC Berkeley, pp 130.
- Dane, L.J. 2014. Soil Microorganisms as Precursors and Mediators of Soil Carbon Stabilization. Ph.D. Thesis, UC Berkeley, pp111.
- Estera, K, 2017. Characterization of Avena barbata exudates. M.S. Thesis. U.C. Berkeley
- Biao Zhu, J.L.M. Gutknecht, D. J. Herman, D.C. Keck, M. K. Firestone, W. Cheng. 2014. Rhizosphere priming effects on soil carbon and nitrogen mineralization. Soil Biol. Biochem. 76:183-192.
- Shi, S., E. Nuccio, DJ Herman, R Rujkers, K Estera, J Li, UN da Rocha, Z He, J Pett-Ridge, EL Brodie, J Zhou, M Firestone. 2015. Successional Trajectories of Rhizosphere Bacterial Communities over Consecutive Seasons. mBio.00746-15.

- Nuccio, E. E., J. Anderson-Furgeson, K.Y. Estera, J. Pett-Ridge, P. de Valpine, E. L. Brodie, M.K. Firestone. 2016. Climate and edaphic controllers influence rhizosphere community assembly for a wild annual grass. Ecology 97:1307-1318.
- Shi, Shengjing, Erin E. Nuccio, Zhou J. Shi, Zhili He, Jizhong Zhou, and Mary K. Firestone. 2016. The interconnected rhizosphere: High network complexity dominates rhizosphere assemblages. Ecol. Letts. 19:926-936.
- Pett-Ridge, Jennifer and Mary K. Firestone, 2017. Using stable isotopes to explore root-microbe-mineral interactions in soil. Rhizosphere. doi.org/10.1016/j.rhisph.2017.04.016