Title: Genomes to Ecosystem Function: Targeting Critical Knowledge Gaps in Soil Methanogenesis and Translation to Updated Global Biogeochemical Models

Authors: Angela M. Oliverio[1,2*] (Angela.Oliverio@colostate.edu), Adrienne B. Narrowe[1+], Mikayla A Borton[1,3], Jorge A. Villa[4], Rebecca A. Daly[1], Pengfei Liu[1], Bridget B. McGivern[1], Michael Shaffer[1], Rory Flynn[1], Christopher S. Miller[4], Trent R. Northen[5], David W. Hoyt[3], William R. Riley[5], Timothy H. Morin[7], Gil Bohrer[8], **Kelly C. Wrighton**[1]

Institutions: 1 Colorado State University, Fort Collins; 2 Yale University, New Haven; 3 Pacific Northwest National Lab, Richland, WA; 4 University of Colorado, Denver; 5 Lawrence Berkeley National Lab, Berkeley, CA; 6 University of Louisiana, Lafayette; 7 State University of New York ESF, Syracuse; 8 The Ohio State University, Columbus

Project Goals:

Despite their relatively small land coverage, wetlands represent the largest source of atmospheric methane (20-40%). However, variations in these wetland emission budgets are high, with over 25% uncertainty. Accurately predicting net methane fluxes from these systems depends on multiple interacting geochemical, ecological, and metabolic constraints that are poorly understood, oversimplified, or missing in global biogeochemical models. The overarching objective of this early career proposal is to identify the biogeochemical and genomic determinants impacting methane production, including the scale and physical distribution over which they operate, along freshwater wetland gradients.

The project goals can be summarized into three main objectives:

1) Characterize wetland variables affecting methanogen distribution and activity.

2) Quantify fine scale oxygen dynamics in soils and the impact on methane emissions.

3) Improve model capability to include a representation for small-scale processes.

Abstract Text:

Freshwater wetlands are important contributors to ecosystem services and mediate global biogeochemical cycles with important contributions to soil carbon sequestration and greenhouse gas (GHG) flux. Toward our goal of understanding controllers on microbial carbon decomposition in soils, we constructed a greenhouse gas (GHG) observatory at Old Woman Creek, a freshwater wetland in the Ameriflux network with some of the highest annual methane fluxes [1]. To track the end products of microbial carbon decomposition we measured methane and carbon dioxide production at cm scale depths in soils and GHG fluxes from soils at meter and wetland scales. We defined the methane and carbon dioxide production hotspots occurring within the top 50 cm of the soil profile and extended these data across multiple months and years. Paired to these data we collected over 600 microbial samples across the wetland including genome-resolved metagenomics with metatranscriptomes, 16S rRNA gene sequencing, and metabolomics (NMR, LC-MS, and FT-ICR MS). During our 5-year sampling period, climatically-driven water level rise in the adjacent Lake Erie flooded the wetland by ~3-6 ft, with this hydrological inundation serving as a natural laboratory for interrogating how climatic

induced shifts in hydrology impact microbial metabolism to regulate soil carbon decomposition and flux.

Essential to our microbial analyses, we constructed a first of its kind genomic catalog of freshwater soil microbial communities, resulting in over 13,000 genomes spanning 64 bacterial and 13 archaeal phyla. To first understand the microbial processes directly regulating soil methane, we genomically inventoried 300 genomes from methanogen and methanotroph lineages, defining a new family and 12 new genera. We coupled this genomic catalog to metatranscriptome data from spatially and temporally distributed samples along with geochemical, metabolomic, and gas flux measurements. Here we provide evidence that contrasts with widely held paradigms about the soil carbon cycle: (1) methanogenesis using methylated organic nitrogen and sulfur substrates, not acetate or hydrogen, best explained soil methane production, (2) soil oxygen concentrations or depth are not effective approximates for the activity of methanogens or methanotrophs in soil cores [2,3], and (3) redox changes caused by sustained flooding, did not alter the methanogenic substrate pools or the metabolic networks governing methane production. Collectively, these results highlight the soil microbial metabolisms influencing the terrestrial methane cycle, thereby offering direction for increased realism in predictive process-oriented models of methane flux in wetland soils.

To better understand the nutrient and carbon metabolic networks that control soil carbon stability and its decomposition into GHGs, we used our extensive metatranscriptome data set to determine coordinated gene expression networks across season and depth and correlated these networks to key ecosystem chemical features. Our analysis revealed cohorts of metabolically active microorganisms and their associated metabolites that persisted across months and even years, impervious to different soil geochemical conditions. For example, we found the same microbial network predicted soil carbon dioxide and methane production, as well as sulfate concentrations. This active soil organic matter decomposing cohort was comprised of soil carbon respiring members that produce substrates for simple carbon utilizing sulfate reducing bacteria, which generate the methanogenic substrates to support active methane cycling microorganisms. The strong co-occurrence between sulfate-reducers and methanogens is of particular interest, as it is often thought that sulfate reduction is a thermodynamic constraint on methane production in wetland soils. Alternatively, we suggest these groups act syntrophically, as the high dissolved organic carbon is not a competitive constraint, and thus sulfate reducers oxidize lactate and ethanol to hydrogen and acetate, which cross-feeds methanogens. We are now using field-designed reactor studies amended with labeled substrate to validate and probe the resiliency of these carbon decomposing metabolic networks.

While we have demonstrated that microbial genome resolution does increase predictions of emergent ecosystem properties like soil methane production, omics data are not easily ingested into terrestrial ecosystem biogeochemistry models (e.g., *ecosys*). These models incorporate data across ecosystem layers including climate, macrofauna, chemistry, and microbes including microbial functional guilds. Therefore, we developed a software application to summarize individual genome-resolve transcriptomic data into functional guilds using DRAM (Distilled Refined Annotation of Metabolism) [4,5], which was developed as part of this early career award. We demonstrated that DRAM identifies trait-level information by leveraging annotation

and metabolism content from microbial genomes and expression patterns and translates this into "functional guilds".

To test if we could predict the observed variability in soil carbon dioxide and methane using only our functional guild data, we built random forest models and evaluated the importance of predictors (e.g., functional guilds) by determining the decrease in prediction accuracy between observations and predictions. For soil methane production, the strongest guild predictors were methanogens, homoacetogens, and fermenters, while aerobic heterotrophs and nitrifiers predicted carbon dioxide production. Lastly, these explanatory models and coordinated gene expression networks informed a conceptual, community-based metabolic model identifying possible abiotic and biotic control points on the interconnected microbial metabolisms that control carbon dioxide and methane production. Our analyses suggest that translating complex microbial data into a 'guild' framework may represent a path to readily incorporate highly dimensional multi-omics into terrestrial ecosystem biogeochemistry models while still maintaining biological integrity.

References

1. Rey-Sanchez AC, TH Morin, KC Stefanik, KC Wrighton, G Bohrer 2018. Determining total emissions and environmental drivers of methane flux in a Lake Erie estuarine marsh. *Ecological Engineering*.114, 7-15

2. Angle JC, TH Morin, LM Solden, GJ Smith, AB Narrowe, MA Borton, RA Daly, DW Hoyt, WR Riley, CS Miller, G Bohrer, KC Wrighton 2017. Methanogenesis in oxygenated soils is an unrecognized driver of wetland methane emissions. *Nature Communications*, (8)1, 1567

3. Smith GJ, JC Angle, LM Solden, MA Borton, TH Morin, RA Daly, MD Johnston, KC Stefanik, R Wolfe, G Bohrer, KC Wrighton 2018. Members of the Methylobacter are inferred to account for the majority of aerobic methane oxidation in oxic soils from a freshwater wetland. *mBio* 9 (6) e00815-18.

4. Shaffer M, MA Borton, BB McGivern, AA Zayed, SL La Rosa, LM Solden, P Liu, AB Narrowe, J Rodriguez-Ramos, B Bolduc, C Gazitua, RA Daly, GJ Smith, DR Vik, PB Pope, MB Sullivan, S Roux, KC Wrighton 2020. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Research*.

5. Rubinstein RL, MA Borton, H Zhou, M Shaffer, DW Hoyt, J Stegen, CS Miller, C Henry, KC Wrighton, R Versteeg. ORT: A workflow linking genome-scale metabolic models with reactive transport codes. BioRxiv https://doi.org/10.1101/2021.03.02.433463

Funding Statement: This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0018022. This program is supported by the U. S. Department of Energy, Office of Science, through the Genomic Science Program, Office of Biological and Environmental Research, under FWP ERKP123.