Title: Soil Moisture Impacts Composition, Activity and Ecosystem Functions of Soil DNA and RNA Viruses

Authors: Ruonan Wu[1] (ruonan.wu@pnnl.gov), Michelle R. Davison[1], William C. Nelson[1], Yuqian Gao[1], Carrie D. Nicora[1], Jason E. Mcdermott[1], Kristin E. Burnum-Johnson[1], **Kirsten S. Hofmockel[1, 2], Janet K. Jansson**[1]

Institutions: [1]Earth and Biological Sciences Directorate, Pacific Northwest National Lab, Richland, WA, USA; [2]Department of Agronomy, Iowa State University, Ames, IA, USA

https://www.pnnl.gov/projects/soil-microbiome-science-focus-area

Project Goals: PNNL's Phenotypic Response of Soil Microbiomes SFA aims to achieve a systems-level understanding of the soil microbiome's phenotypic response to changing moisture. We perform multi-scale examinations of molecular and ecological interactions occurring within and between members of microbial consortia during organic carbon decomposition, using chitin as a model compound. Integrated experiments are designed to confront spatial challenges and inter-kingdom interactions among bacteria, fungi viruses and plants that regulate community functions. These data are used to parametrize individual- and population-based models for predicting interspecies and inter-kingdom interactions. Predictions are tested in lab and field experiments to reveal individual and community microbial phenotypes. Data is captured and shared through an optimized data management pipeline. Knowledge gained will provide fundamental understanding of how soil microbes interact to decompose organic carbon and enable prediction of how biochemical reaction networks shift in response to changing moisture regimes.

Soil is known to harbor a large diversity of viruses, but the majority are uncharacterized and how the soil environment impacts the composition, activity and ecosystem functions of soil viruses is unknown. Here, we screened for viral sequences in some of the largest soil metagenomes produced to date (>1 Tb each) from a range of grassland soils with different historical precipitation patterns - from arid soils in eastern Washington to wet soils in Iowa and intermediate soils in Kansas. Due to the unprecedented size of the metagenomes, we collaborated with NERSC to assemble them[1]. Screening the assemblies resulted in a total of 2,631 viral contigs including 14 complete circular viral genomes[2], which were grouped into 214 clusters based on the protein sharing matrix and tetranucleotide frequency. A comparison of the viral communities across each site revealed that soil with lower historical moisture harbored significantly higher viral diversity and abundance, while also displaying less evidence of virus-host interactions (through CRISPR-Cas spacers). These seemly contradictory findings may reflect a predominance of viruses with lysogenic strategies in drier soils. Across the diverse grassland DNA virosphere, 45 of the viral clusters were found in all three grasslands. These viral 'common clusters' also featured high abundances, targeted the dominant bacterial taxa, contained putative viral generalists and were more frequently targeted by host CRISPR-Cas spacers. We hypothesize that these 'common clusters' contain viruses that are well adapted to grassland soils and represent the potential viral keystone species in these soils. In addition, auxiliary metabolic genes (AMGs) involved in 18 pathways were identified, with many detected on the complete viral genomes we assembled from the grassland metagenomes, suggesting potential novel viral contributions to carbon metabolism and energy acquisition in soil.

To further investigate the response of DNA and RNA viruses to changes in soil moisture, we collected samples from the Kansas grassland with a historical intermediate soil moisture profile. Three replicate samples from two field locations were either wet to saturation or air-dried for 15 days at 21°C to represent experimental wet and dry soil treatments, respectively. Paired expression data

(metatranscriptomes and metaproteomes) were mapped back to the assembled viral contigs to identify active DNA viruses and to determine their phenotypic responses to soil moisture extremes. The vast majority of actively transcribed DNA viruses were bacteriophage, but some were assigned to eukaryotic hosts, mainly insects[3]. We observed that higher soil moisture increased the transcription of a subset of DNA viruses. In contrast, there was a lower overall level of transcription in drier soil, but across a larger range of DNA viruses. A higher percentage of non-coding RNAs and more transcripts of lysogenic markers (i.e. integrases and excisionases) were also detected in dry soil, further supporting a higher prevalence of lysogenic viruses in more arid soils. We also detected peptides encoded by viral early and late genes that are known to be upregulated during an active viral infection. To our knowledge, this is the first use of bulk soil metaproteome data to detect viral peptides for identification of potentially active viruses. The metatranscriptome data was also used to assemble RNA viruses and illuminated a high sequence diversity in the soil RNA virosphere, revealing a high abundance of *Reoviridae* sequences and a highly diverse population of *Leviviridae*. A hallmark of *Leviviridae* is their lytic lifestyle and thus their higher abundance in wet soil suggests that there was greater potential for host lysis under high moisture conditions.

In summary, our combination of deep metagenomic sequencing with a multi-omics evaluation of viral gene expression provided new details about the influence of changing soil moisture, historical or experimental, on the composition, activity and potential ecosystem functions of soil viruses.

References:

1 Nelson, W.C., Anderson, L.N., Wu, R., McDermott, J.E., Bell, S.L., Jumpponen, A., Fansler, S.J., Tyrrell, K.J., Farris, Y., Hofmockel, K.S. and Jansson, J.K.. Terabase metagenome sequencing of grassland soil microbiomes. *Microbiology Resource Announcements* **9** (2020).

2 Wu, R., Davison, M.R., Nelson, W.C., Graham, E.B., Fansler, S.J., Farris, Y., Bell, S.L., Godinez, I., Mcdermott, J.E., Hofmockel, K.S. and Jansson, J.K. DNA viral diversity, abundance and functional potential vary across grassland soils with a range of historical moisture regimes. (2021, in preparation).

3 Wu, R., Davison, M.R., Gao, Y., Nicora, C.D., Mcdermott, J.E., Burnum-Johnson, K.E., Hofmockel, K.S. and Jansson, J.K. Moisture modulates soil reservoirs of active DNA and RNA viruses. (2021, in review)

***Funding statement*:** This work was supported by the Department of Energy (DOE) Office of Biological and Environmental Research (BER) and is a contribution of the Scientific Focus Area "Phenotypic response of the soil microbiome to environmental perturbations". PNNL is operated for the DOE by Battelle Memorial Institute under Contract DE-AC05-76RL01830. Portions of this work was performed in the W. R. Wiley Environmental Molecular Sciences Laboratory (EMSL) (grid.436923.9), a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research and located at Pacific Northwest National Laboratory (PNNL). The metaproteomic work was funded by the U.S. Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research (BER), Early Career Research Program to KEBJ.