Viral ecology through time and along a permafrost thaw gradient

Christine L. Sun*[1,2,3 †] (sun.2508@osu.edu), Akbar Adjie Pratama[1,2,3 †], Consuelo Gazitua[4], Ahmed Zayed[1,2,3], Dylan Cronin[1,2,3], Lindsey Solden[1], Benjamin Bolduc[1,2,3], the IsoGenie Project Field Teams 2010-2017, the IsoGenie Project Coordinators, Virginia Rich[1,2,3], **Matthew B. Sullivan**[1,2,3]

[1]The Ohio State University, Columbus, Ohio, USA; [2]Center of Microbiome Science Columbus, Ohio, USA; [3]EMERGE Biology Integration Institute, Columbus, Ohio, USA; [4]Viromica Consulting, Santiago, Chile. [†] Authors contributed equally to this work.

Project goals: The overarching goal of this project is to establish ecological paradigms for how viruses alter soil microbiomes and nutrient cycles by developing foundational (eco)systems biology approaches for soil viruses. Within this overall project, we examined viruses from metagenomic datasets to investigate viral ecology across a nearly decadal time series. Such work is critical for establishing baseline ecological understanding of viruses in soils, as well as elucidating the role of viruses in terrestrial carbon and nutrient cycling.

Permafrost is thawing due to elevated temperatures resulting from climate change. It is important to understand how thawing permafrost will impact the release of greenhouse gasses since permafrost accounts for 30%-50% of global soil carbon (C). Microorganisms play a critical role in the terrestrial C cycle, but viruses are less well-studied in soils. However, in marine systems, viruses are known to impact C cycling by controlling host microbial communities via predation, transferring genes between hosts, and metabolically reprogramming host cells via encoding auxiliary metabolic genes (AMGs). Here, we examined viruses identified in 379 bulk soil metagenomes derived along a permafrost thaw gradient (palsa, bog and fen) in Stordalen Mire, Sweden, over eight years (2010-2017), and use these data to establish baseline ecological understanding. In total, we identified 8,597 unique viral operational taxonomic units (vOTUs; ≥ 5 kb contigs dereplicated at 95% average nucleotide identity and 80% coverage), which gene-sharing network analyses suggested represent 1,609 genera, of which 617 are novel (compared to NCBI Viral RefSeq). Ecologically, many vOTUs were shared across sites (e.g., 50% of bog vOTUs were shared with at least one other site), but ordination separated viral communities based on their origin along the thaw gradient. Alpha diversity differed within each site, and increased with depth. While some vOTUs were found in each site through the entire time series, the vast majority of OTUs were ephemeral as they were only detected in a single sample. Hosts for these viruses were *in silico* predicted, with the benefit of co-sampled metagenome-assembled genomes (MAGs), and this revealed 830 MAGs as hosts for

992 vOTUs. Of these MAGs, 790 (95%) have genome encoded functions that include degradation of specific carbon compounds (such as cellulose, fructose, and xylose), implicating viruses in C cycling via infection. In addition, we identified 1,349 instances of 81 unique AMGs in vOTUs from a diversity of metabolic genes, including those involved in central C metabolism and C degradation (e.g. glycoside hydrolase families). Together, these results provide baseline ecological patterns, as well as evidence that soil viruses may impact permafrost C cycling through diverse means.

Funding statement: This work was supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program under Award Numbers DE-SC0020173 (to MBS), DE-SC0004632, DE-SC0010580, and DE-SC0016440 (to VIR), as well as the National Science Foundation #1758974 (to MBS) and #2022070 (to VIR).