Pacific Northwest
NATIONAL LABORATORY

PNNL-30560

Uncovering novel RNA viruses in permafrost

October 2020

Michelle R Davison
Ruonan Wu
Vincent G Danna
Iobani G Godinez

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <https://www.ntis.gov/about>
Online ordering: http://www.ntis.gov

PNNL-30560

Uncovering novel RNA viruses in permafrost

October 2020

Michelle R Davison
Ruonan Wu
Vincent G Danna
Iobani G Godinez

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Uncovering novel RNA viruses in permafrost

Project Summary: This study aims to unveil the taxonomic diversity and potential ecological functions of RNA viruses in permafrost

Introduction and Project Description:

Arctic soils (including permafrost) contain an untapped reservoir of microbial life, including viruses, that are trapped in a frozen state of low metabolic activity[1]. However, as permafrost thaws, they can be revived and pose potential risks to human health and ecosystem stability. The current COVID-19 pandemic has highlighted the urgency of identifying potential pathogen reservoirs which may be activated under the impacts of climate change, as well as understanding the role of RNA viruses in dynamic environments. Here, we applied a new computational workflow (Wu et al. submitted) to identify and analyze RNA viruses from *de-novo* assemblies representing a total of 33 metatranscriptomes collected from four parallel transects across a range of fluctuating environmental gradients[2]. This study provides the first characterization of RNA viral diversity within a permafrost ecosystem.

Results and Accomplishments

A total of 22,408 RNA viral contigs were identified which contained 228 unique RdRp sequences (99% amino acid identity) assigned to 22 phylogenetic groups (Fig. 1), including enveloped, non-enveloped, naked ribonucleoprotein-complex viruses, and representatives of dsRNA, (+)ssRNA, and (-)ssRNA clades (Fig. 2a). We observed a high sequence redundancy in the phylogenetic marker, RdRp, in comparison to a recent study on a grassland soil microcosm[3], suggesting fewer RNA viral phylogenetic clades in permafrost. However, eight phylogenetic clades (e.g. *Mitoviridae* and *Hypoviridae*) were identified in the permafrost samples but were absent in grassland soils, highlighting the unique assemblage of RNA viral communities in permafrost.

Several additional significant findings are highlighted as follows: 1) Canonical Correspondence Analysis (CCA) revealed viral clades that were significantly correlated with some environmental measurements, such as proportion of sand, Fe(II), and chloride ($p < 0.05$, Fig. 2b), suggesting that these variables have an impact on RNA viruses. 2) Some members of the eukaryotic community were also significantly correlated to the permafrost RNA viral community (i.e. *Metazoa*, and *Amoebazoa*, $p < 0.05$, Fig. 2c/3a). 3) Putative host-virus relationships were assigned by annotation using the Virus-Host DB[4] and revealed a complex web of infectivity (Fig. 3b), suggesting that RNA viruses may have a widespread influence on permafrost microbial community dynamics and function. 4) We also identified several potential auxillary metabolic genes (AMGs) in high abundance in permafrost RNA viruses, suggesting they play a role in shaping host dynamics and metabolic functions (Fig. 4). For example, one such AMG, encoding a polygalacturonase 1 enzyme known to be involved in degradation of plant and fungi cell walls, was previously implicated as a pathogenicity factor in both fungal and bacterial species.

In summary, these results showcase the capability of our computational workflow for detecting RNA viruses in a complex system. As a result, we were the first to characterizing the permafrost RNA virome and to identify environmental factors that influence the RNA virome community structure. Finally, we provide biological insights into the ecological importance of RNA viruses in permafrost soil microbial population dynamics and metabolism.

Manuscripts:

- Ruonan Wu, Vincent Danna, Iobani Godinez, James Stegen, Eric Bottos, Janet K. Jansson and Michelle Davison. *Uncovering novel RNA viruses in permafrost,* in preparation

Figures:

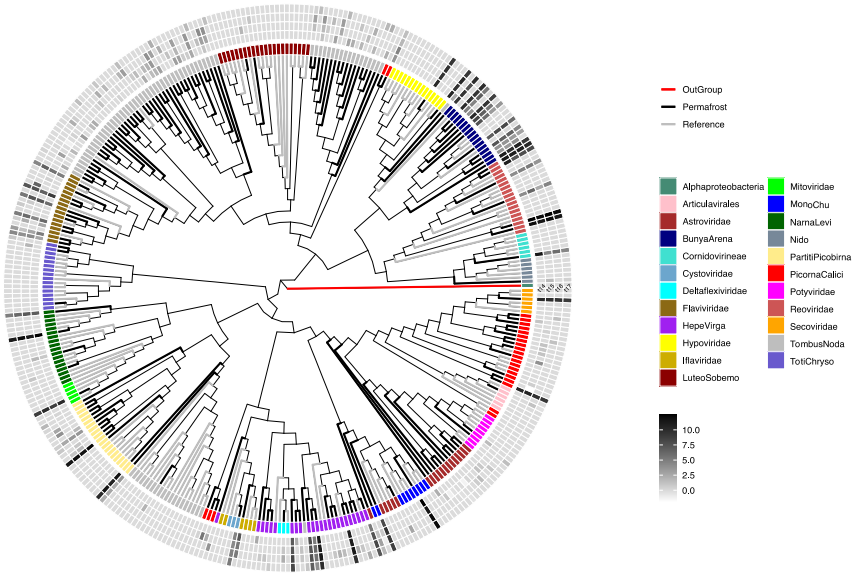

Fig. 1: Rooted phylogeny of RNA viruses based on RdRp alignments. The outgroup branch (DNA-directed RNA polymerase *Alphaproteobacteria* (WP_012231479.1) is highlighted in red. The branches representing sequenced RNA viruses from NCBI viral databases and recent publications[2,4] are shown in grey, while sequences from this study are shown in black. Identified RNA viral phylogenetic clades are shown in the colored inner heatmap ring, while relative abundances normalized across transects [t14, t15, t16, t17] are shown in the outermost rings in greyscale.

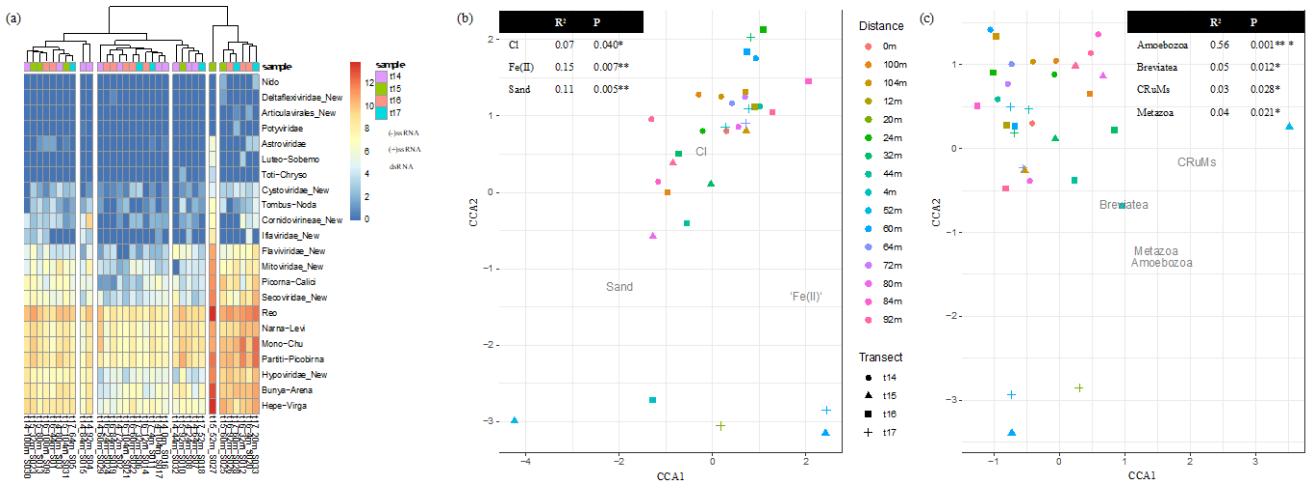

Fig. 2: Permafrost RNA viral community is at least partly shaped by environmental factors and eukaryotic populations. a) RNA viral community composition is shown across transect samples (t14, t15, t16, t17)[2]. Log transformed abundances are visualized as a heat map with low abundances in blue and high in red. RNA viral clades are colored coded by genome type (i.e ssRNA, etc.) b) A CCA plot of RNA viral community composition across transects with a range of environmental parameters[2] revealed three parameters to be statistically significant: proportion of sand, Fe(II) [ug/g], and chloride [ug/g]. c) A CCA plot of RNA community composition with eukaryotic hosts revealed four groups to be statistically significant, *Metazoa*, *Amoebozoa*, *Breviatea*, and *CRuMs*. R^2 and P-values are as shown.

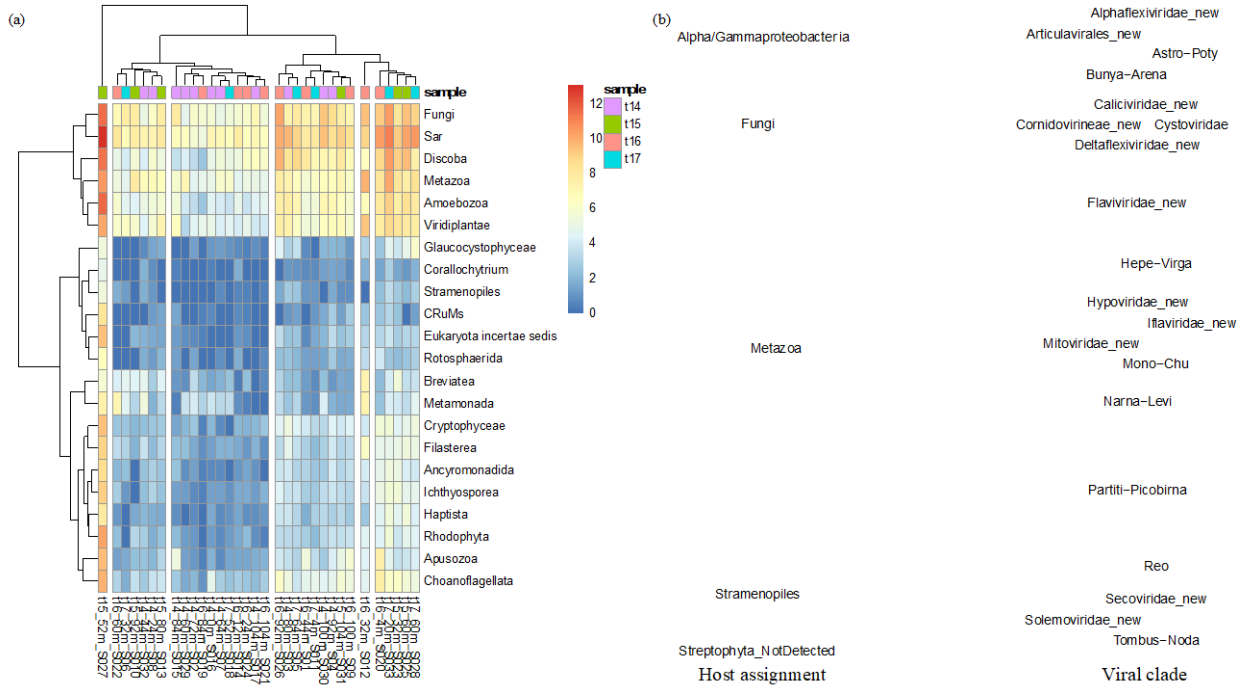

Fig. 3: Eukaryotic community composition and the host assignment of the RNA viral clades detected from permafrost. a) Eukaryotic community composition is shown across transect sampling (t14, t15, t16, t17)[2]. Log transformed abundances are visualized as a heat map with low in blue and high in red. b) Host-viral assignments are shown in an alluvial plot, with host on the right, and RNA viral clade on the right, and the pairings are colored by the host Phyla. The viral clades indicated by "_new" are unique to permafrost as compared to grassland soils[3]. We detected one viral clade assigned to *Streptophyta* which was not detected in our permafrost metatranscriptome and labeled such as "_NotDetected".

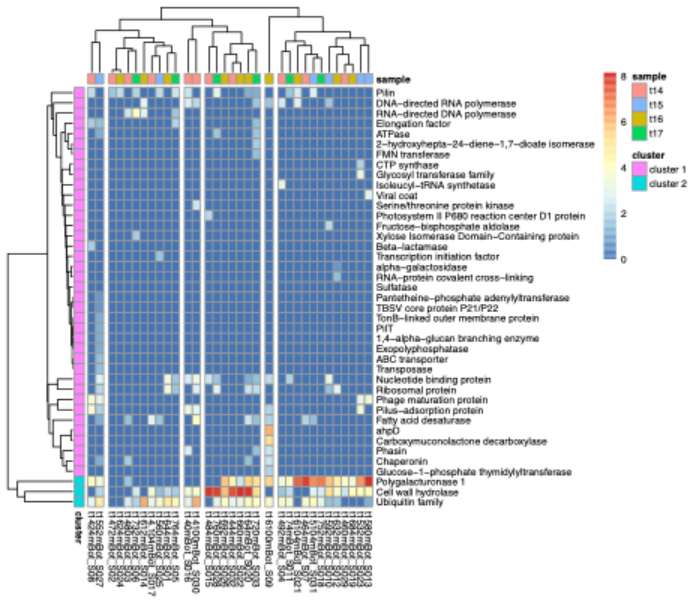

Fig. 4: Annotated auxillary metabolic genes (AMGs) from identified RNA viral contigs. AMG abundances are shown across transect samples (t14, t15, t16, t17)[2]. Log transformed abundances are visualized as a heat map with low abundances in blue and high in red. The three most abundant AMGs are grouped into cluster 2, including AMGS encoding a polygalacturonase 1 enzyme, a cell wall hydrolase, and a protein in the ubiquitin family.

References:

1 Adriaenssens, E. M. *et al.* Environmental drivers of viral community composition in Antarctic soils identified by viromics. *Microbiome* **5**, 83 (2017).

2 Bottos, E. M. *et al.* Dispersal limitation and thermodynamic constraints govern spatial structure of permafrost microbial communities. *FEMS microbiology ecology* **94**, fiy110 (2018).

3 Starr, E. P., Nuccio, E. E., Pett-Ridge, J., Banfield, J. F. & Firestone, M. K. Metatranscriptomic reconstruction reveals RNA viruses with the potential to shape carbon cycling in soil. *Proceedings of the National Academy of Sciences* **116**, 25900-25908 (2019).

4 Mihara, T. *et al.* Linking virus genomes with host taxonomy. *Viruses* **8**, 66 (2016).

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

www.pnnl.gov