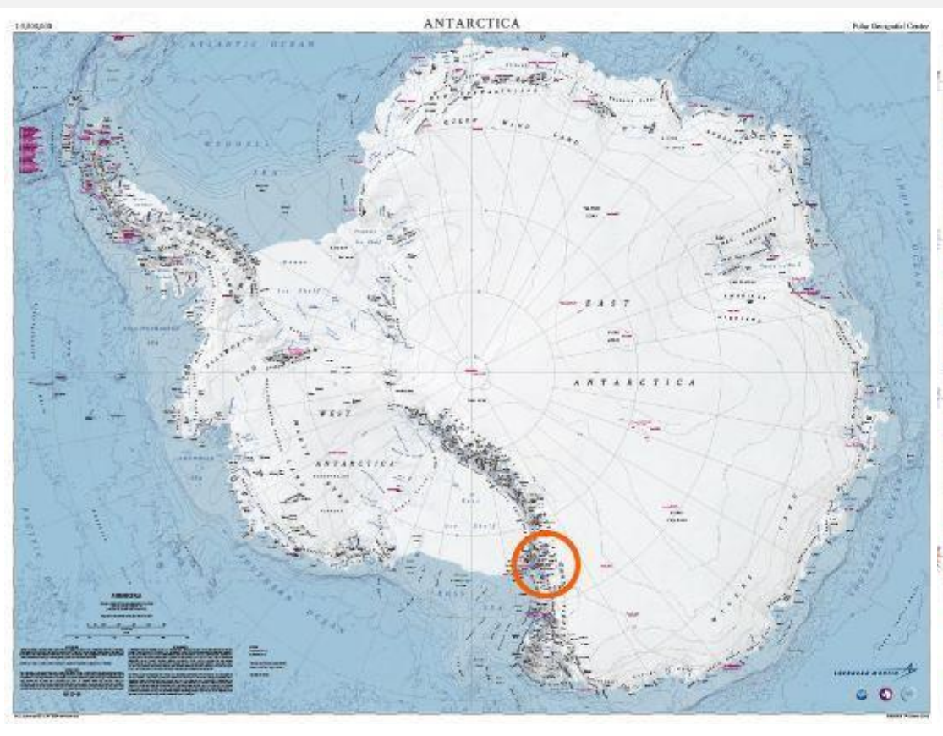
Uncovering the Viral Ecology of Antarctic Soil Microbial Communities

John Henry Lotz-McMillen, Ruby Gilmore, Dr. Shauna Bennett, Dr. Gareth Trubl, Dr. Sarah Stewart Johnson

Georgetown University

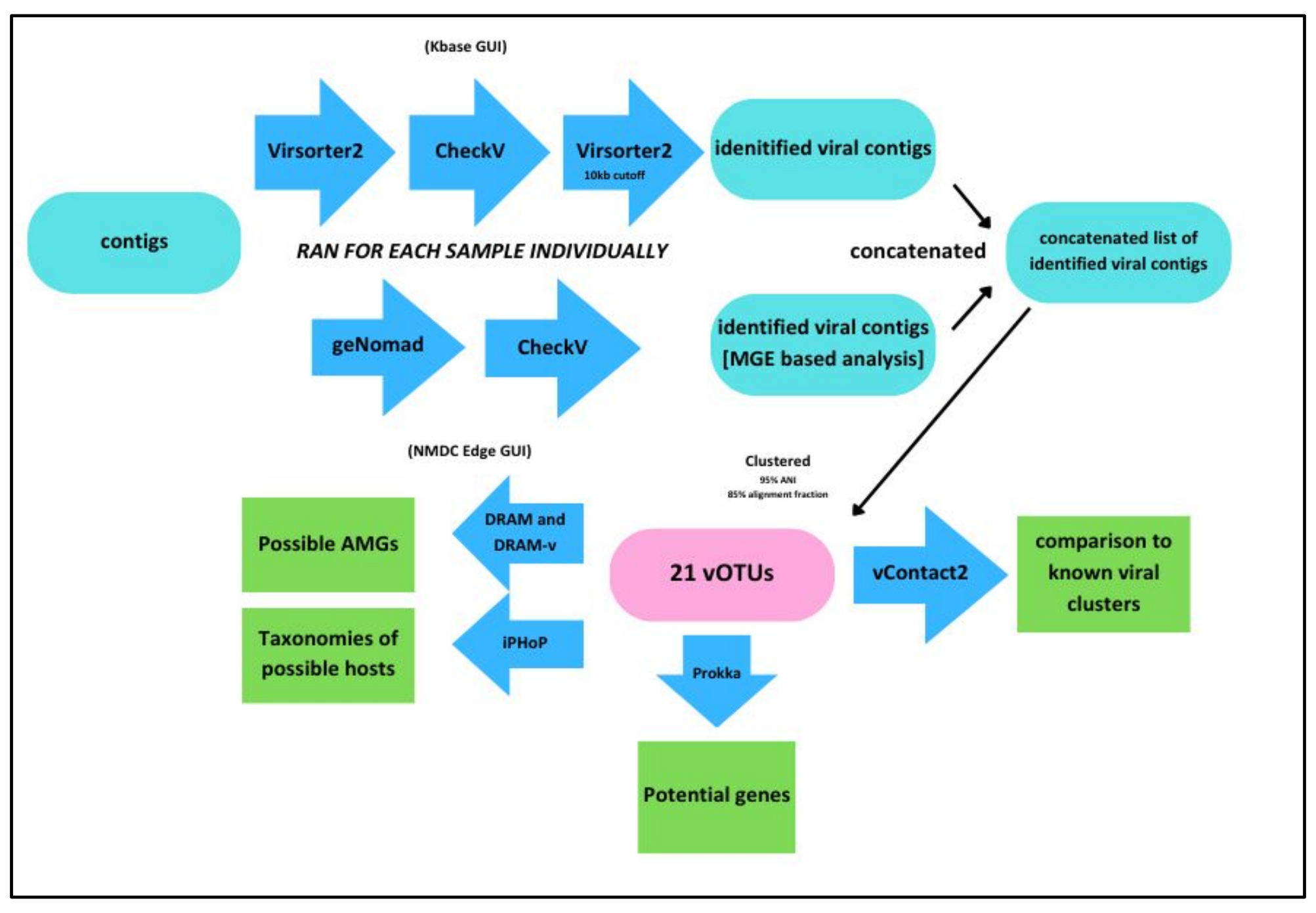
BACKGROUND

- Studying microbial communities in extreme environments offers a unique opportunity to understand how microbial life can persist under harsh conditions.
- Viruses play remarkable but understudied roles in soil ecosystems, managing microbial populations and nutrient cycling while also promoting horizontal gene transfer.
- Research on Antarctic soil viruses is limited given the technological difficulties of distilling viral genetic material from the heterogeneous mixture of inorganic matter and organic biomass in soils.
- Research on Antarctic soil viruses suggests higher levels of diversity at lower altitudes and pH, high levels of heterogeneity independent of geographical proximity, and tailed bacteriophage dominance across studied soils.[2]



- Antarctic soil samples were taken from Taylor Valley in the McMurdo Dry Valleys in December 2016.
- Samples were taken at the surface and at depths of 2cm, 4cm, 6cm, 8cm, and 10cm.

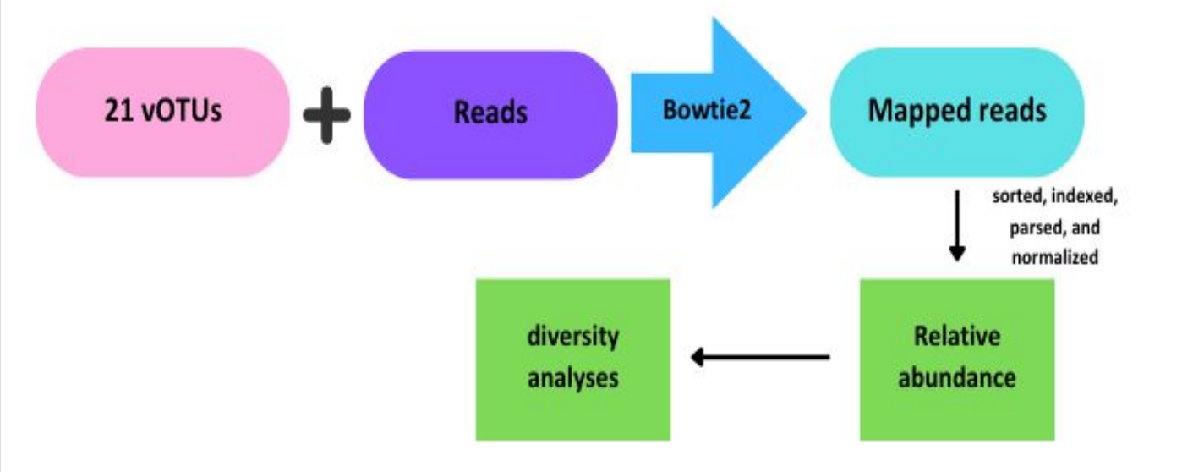
MATERIALS AND METHODS



Workflow for vOTU assembly, vOTU annotation, and taxonomic classification

Sequencing and vOTU Assembly

- 6 soil metagenomic samples were shotgun sequenced.
- VirSorter2[3] and geNomad[4] were used to identify 30 viral contigs ≥10 kb.
- The viral contigs were clustered into a non-redundant set of 21 viral populations (vOTU).



Workflow for read mapping, relative abundance analysis, and diversity analyses

vOTU Analysis

- These vOTUs were functionally annotated with DRAM and DRAM-v[5] and taxonomically classified vContact2[6] and geNomad.
- Hosts were predicted for these viruses using iPHoP.[7]
- Reads were mapped onto the reference genome of vOTUs to identify the relative abundance of each vOTU in each sample.

RESULTS

- Richness was suggested to increase with depth.
- Evenness and diversity is highest at 10cm in depth, decreasing from 4cm to 6cm and then increasing with depth.
- Diversity was suggested to be relatively low across all samples.
- The 6cm and 8cm samples were suggested to be similar in vOTU composition, and the 10cm sample was dissimilar to the other samples in the depth series.
- Nine of the 21 vOTUs were linked to gammaproteobacteria hosts, and only 3 phylum were represented in host prediction (Proteobacteria, Actinobacteriota, and Desulfobacterota B).
- vOTUs were identified as unclassified dsDNA bacteriophages.
- The vOTUs were suggested to belong to the *Caudoviricetes* of tailed bacteriophages.
- DRAM-v identified 1 AMG in vOTU from the 8cm depth sample, transporter gene ABC.PE.P1 involved in the peptides/nickel transport system.

RESULTS

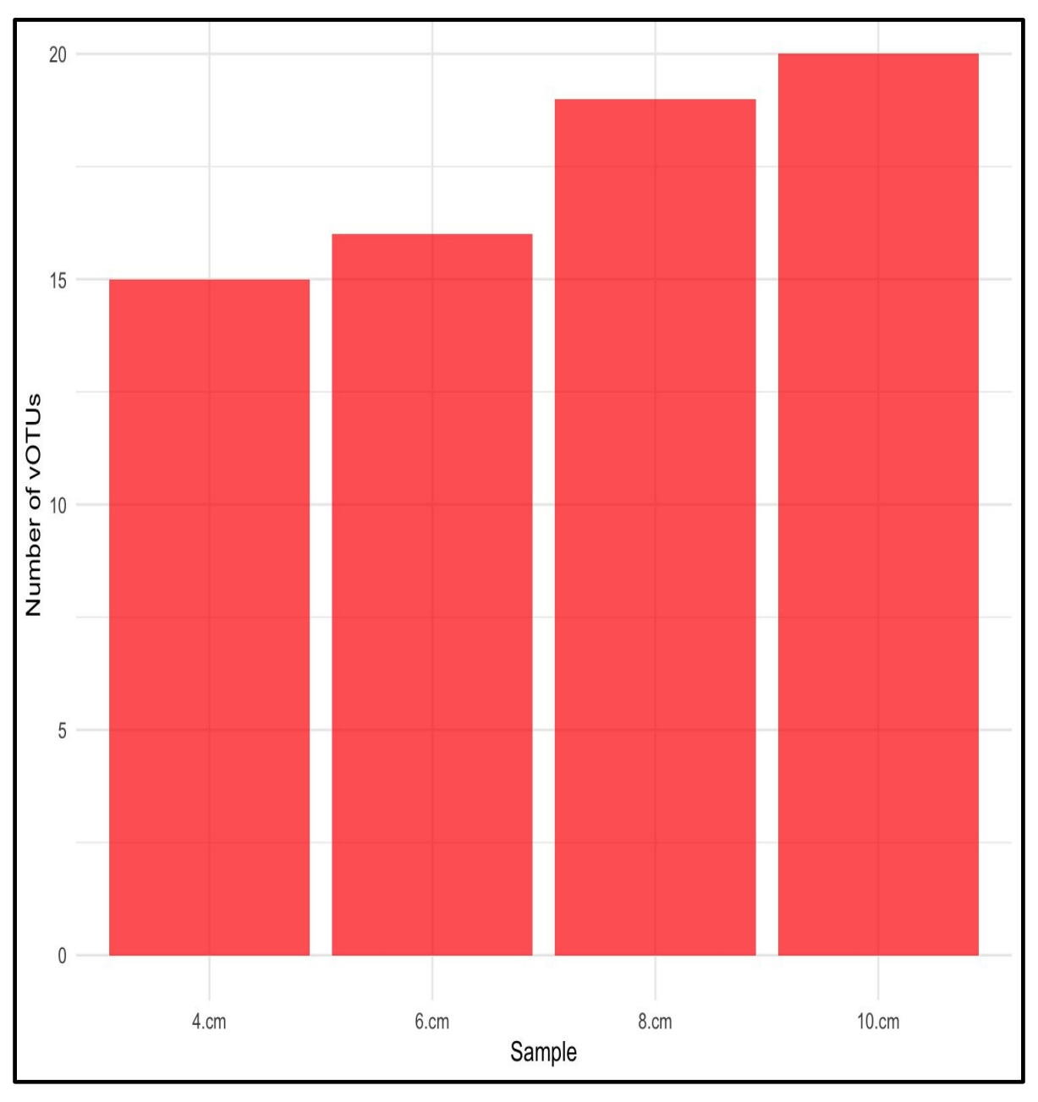


Figure 1. vOTU richness was found using relative abundance table.

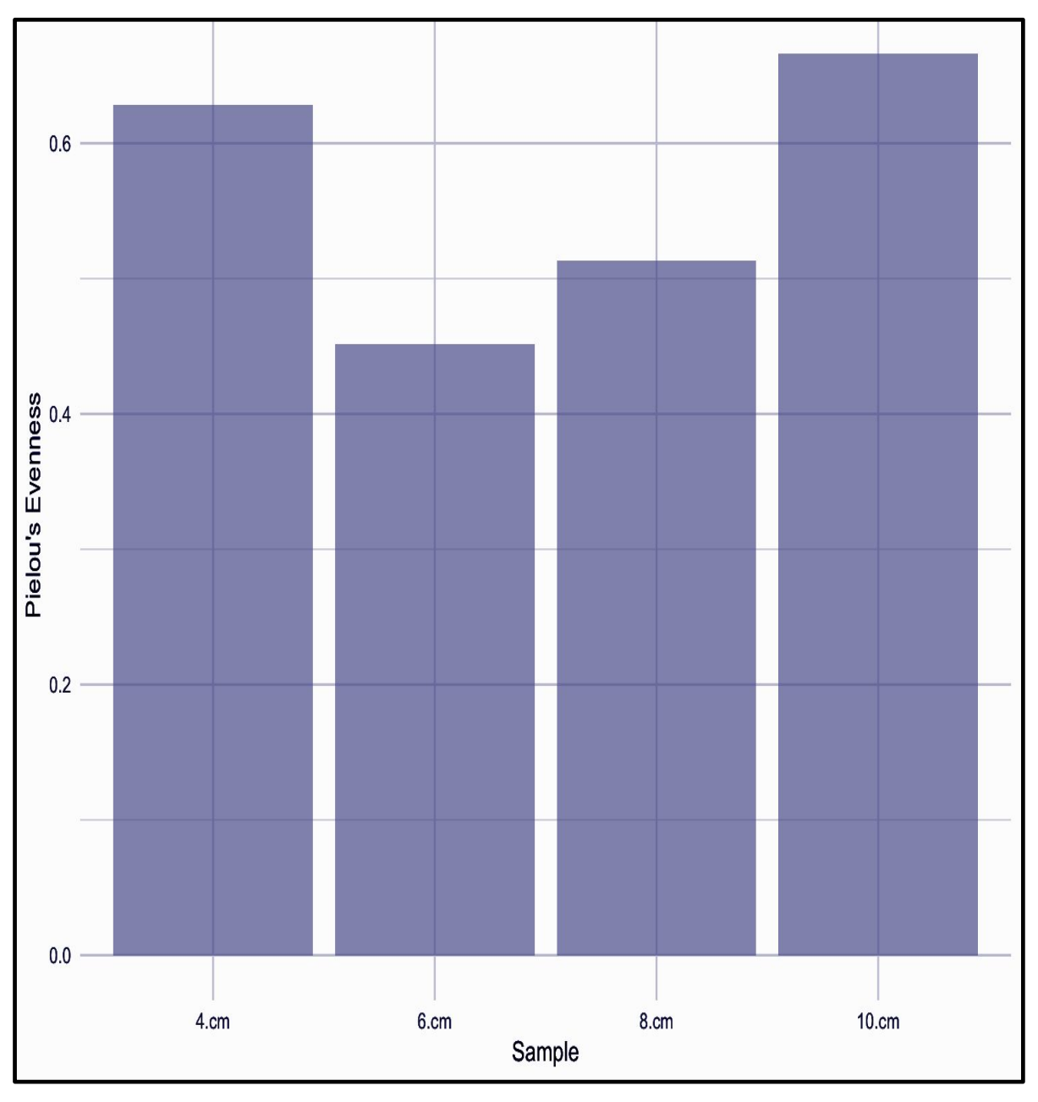


Figure 2. Pielou's evenness was found using relative abundance table.

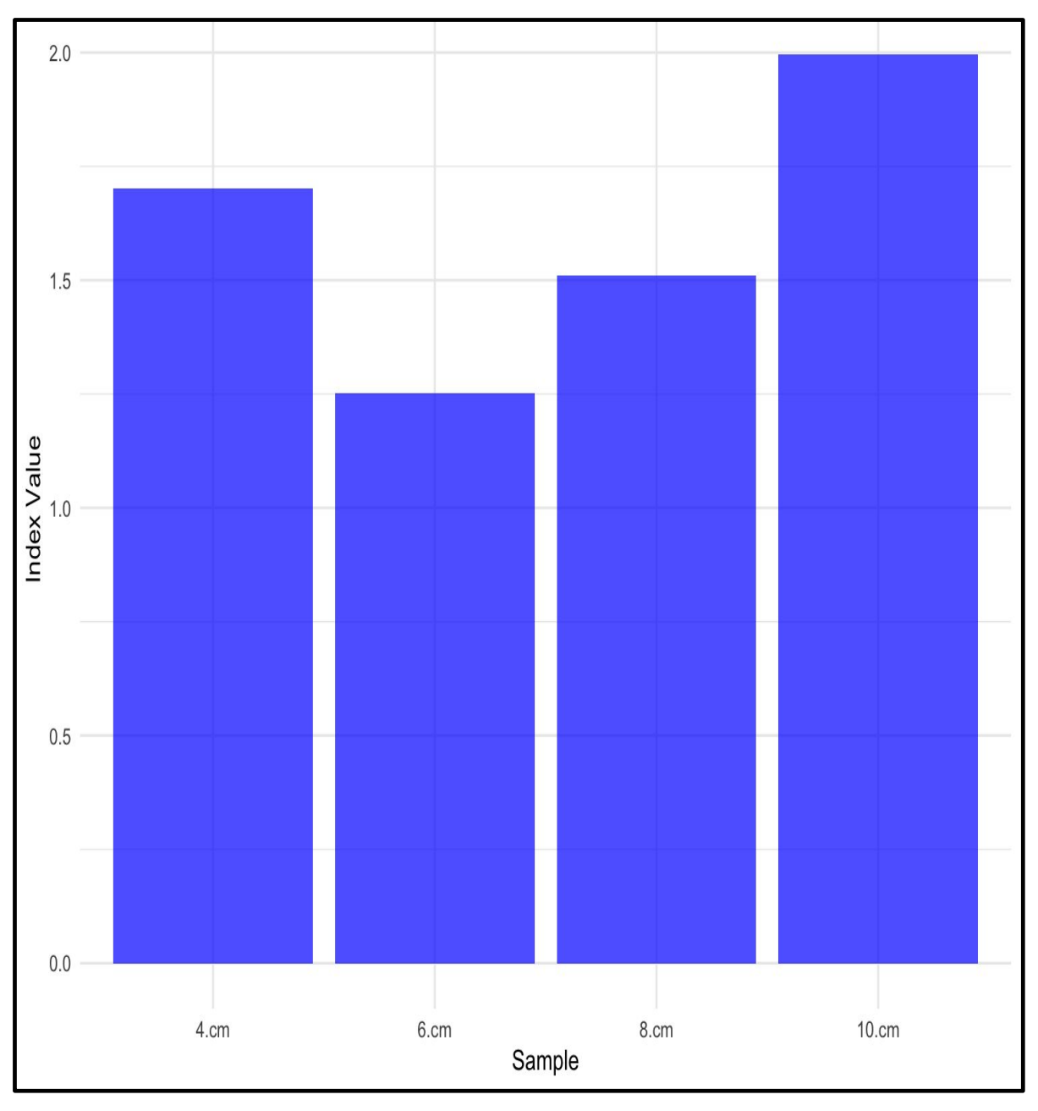


Figure 3. Shannon Diversity was found using relative abundance table.

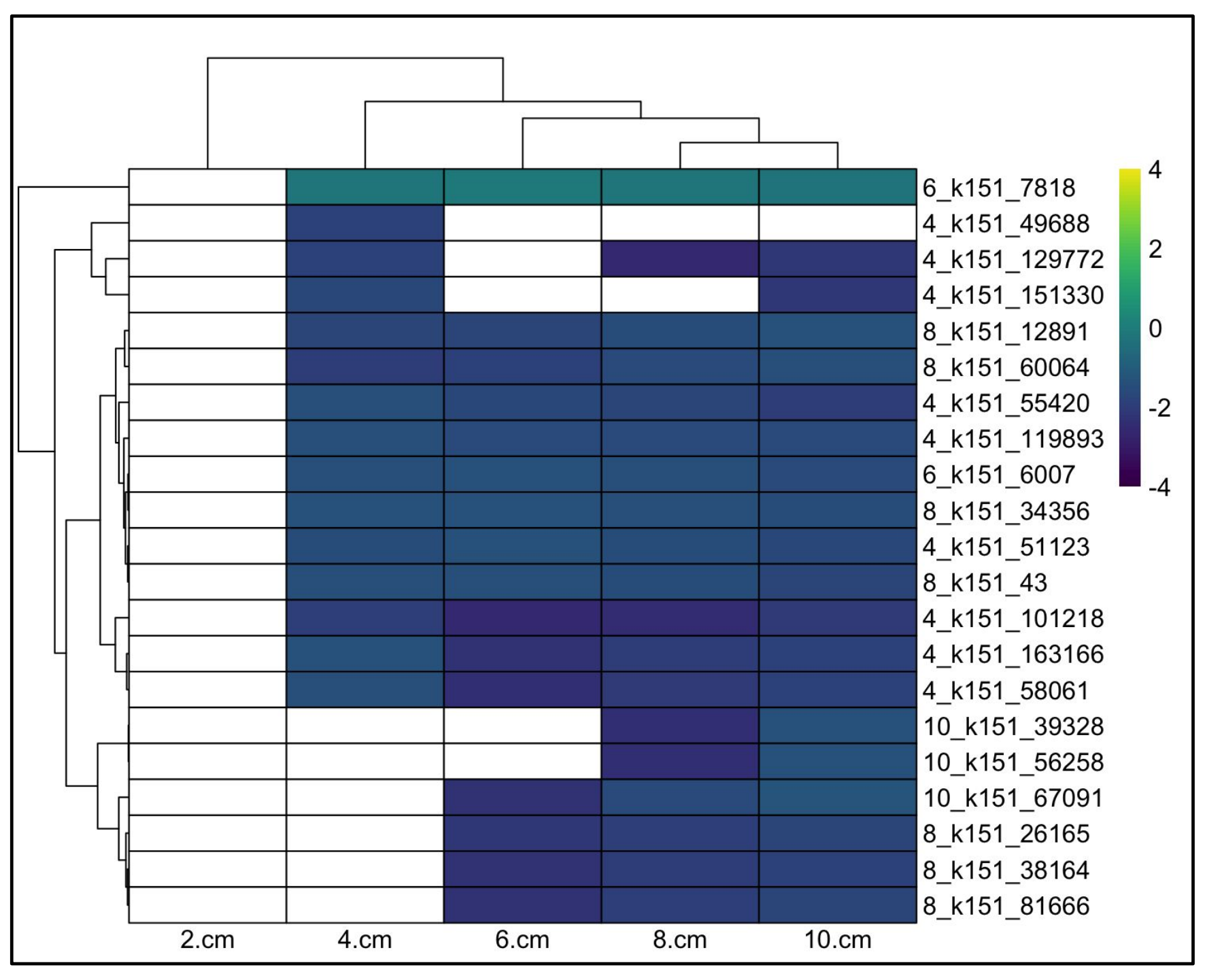


Figure 4. Heatmap of relative abundances. Log 10 transformation of the normalized relative abundance.

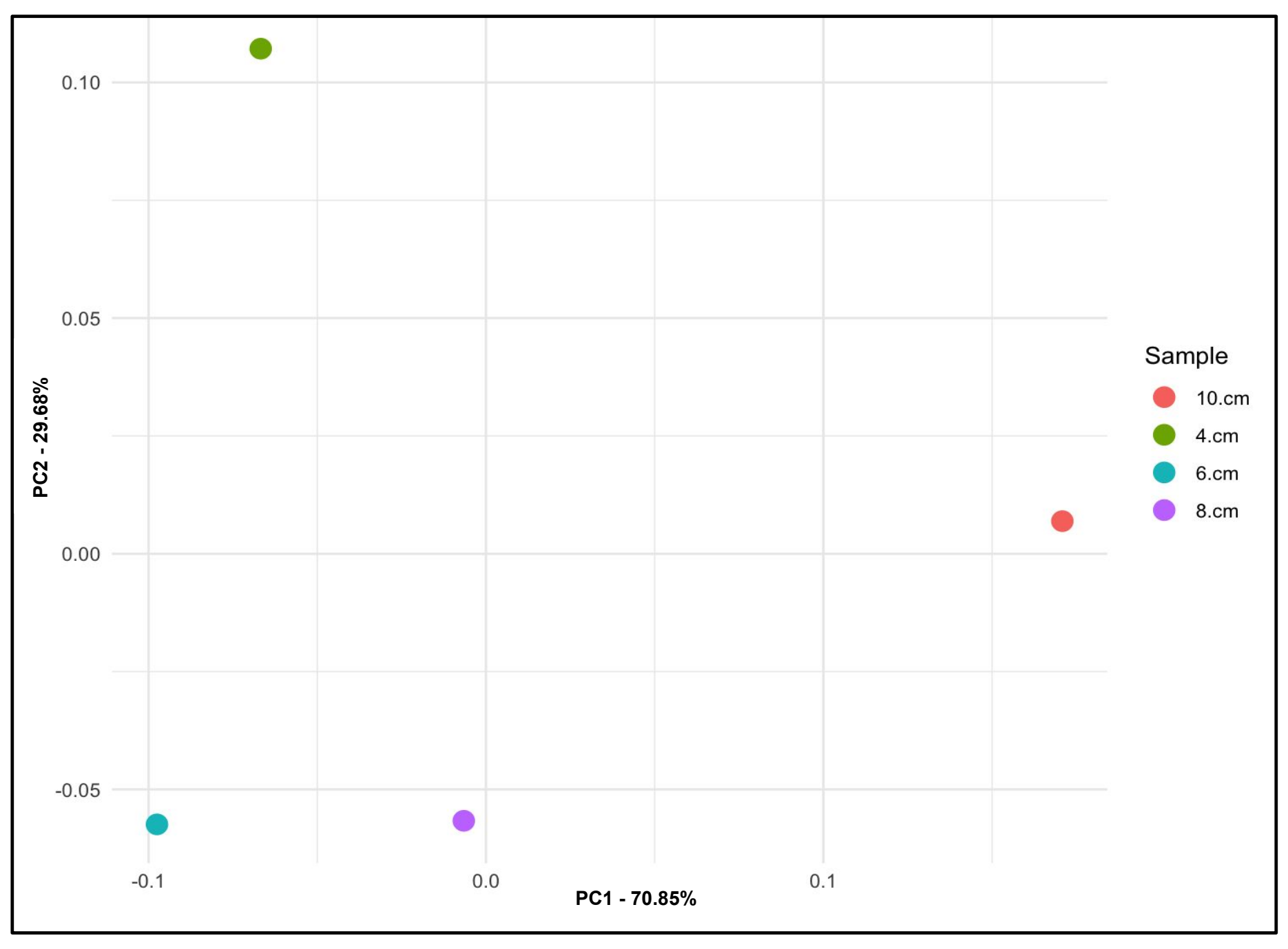


Figure 5. Bray-Curtis PCoA plot that excludes the 0 cm and 2 cm samples.

Figure 6. iPHoP Host Identification. Nine linked to gammaproteobacteria. Five vOTUs did not have predicted hosts. Two vOTUs had two different predicted host classes. Three phyla were represented, Proteobacteria, Actinobacteriota, and Desulfobacterota B (from 1 vOTU)). iPHoP host matches had a confidence score ≥88.

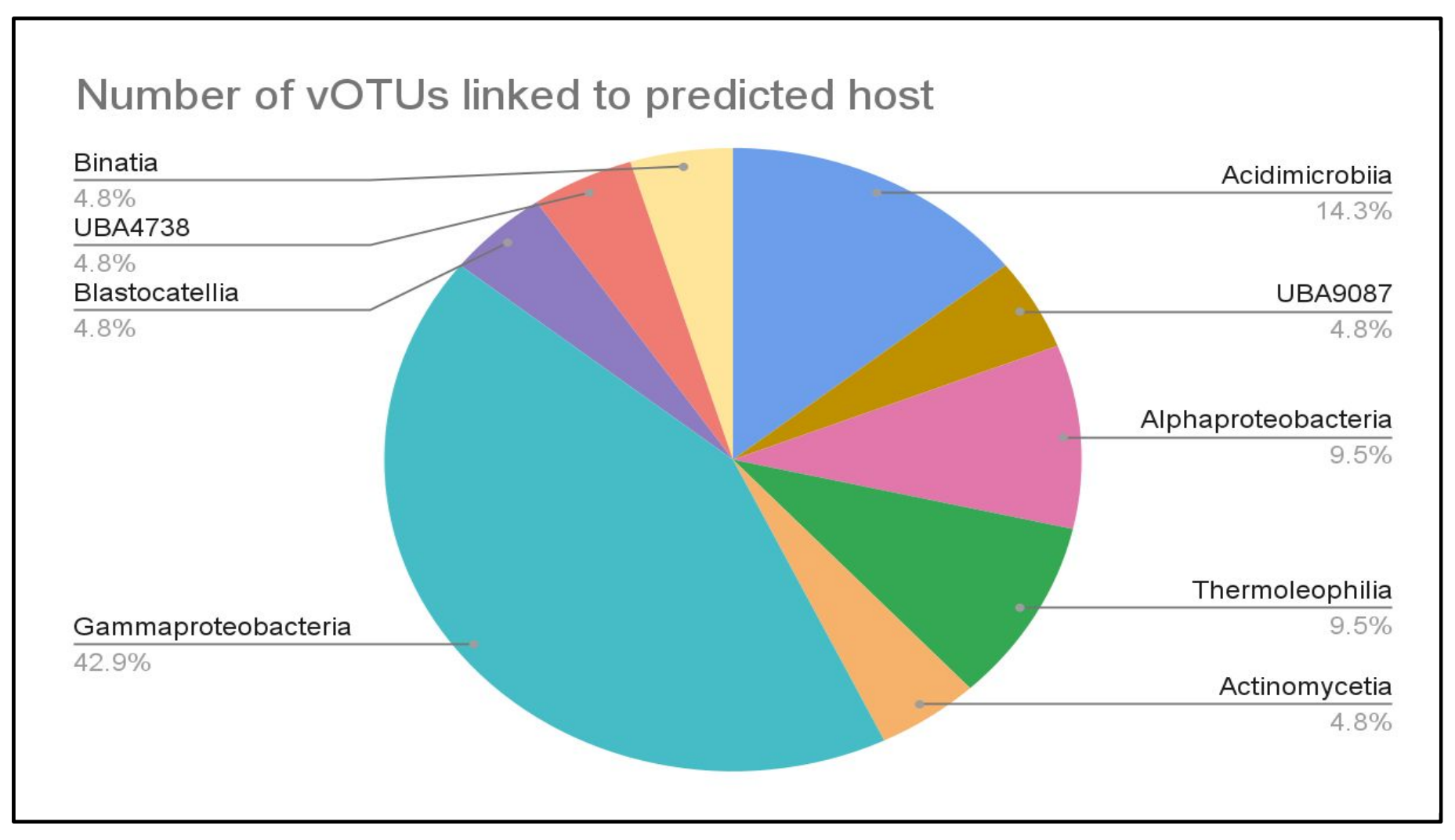


Table 1. vOTU contig length, taxonomic classifications, and gene annotations generated by DRAM-v, IMG/VR BLAST[9], and Prokka.

vOTU (contig ID)	Contig Length, bp	Virus taxonomy (via geNomad or IMG/VR BLAST)	BLAST hits	Annotation
4_k151_101218	10,520	Caudoviricetes	112	[Prokka] DNA gyrase subunit A; DNA gyrase subunit B; DNA replication and repair protein RecF; Chromosomal replication initiator protein DnaA; Beta sliding clamp
4_k151_119893	23,322	Caudoviricetes	163	[Prokka] Tyrosine recombinase XerC
4_k151_129772	10,028	Caudoviricetes	16	[Prokka] tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG; Ribosomal RNA small subunit methyltransferase G; Leucine--tRNA ligase; Chromosome-partitioning protein ParB; Chromosome partitioning protein ParA; IS1595 family transposase ISMtsp22
4_k151_151330	16,344	Caudoviricetes	13	[Prokka] ATP-dependent RecD-like DNA helicase
4_k151_163166	11,977	Caudoviricetes	7	[Prokka] acetyltransferase; Myo-inositol 2-dehydrogenase; transaminase; UDP-N-acetylglucosamine 4-epimerase; UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase; Transcription antitermination protein RfaH; Lipid III flippase
4_k151_49688	11,864	Caudoviricetes	0	[Prokka] DNA-invertase hin
4_k151_51123	18,019	Caudoviricetes	7	[Prokka] Lon protease 2; Sodium-dependent dicarboxylate transporter SdcS; Sodium-dependent dicarboxylate transporter SdcS;
4_k151_55420	16,097	Caudoviricetes	2	[Prokka] carboxylate synthase; Poly(3-hydroxyalkanoate) polymerase; putative diacylglycerol O-acyltransferase tgs1
4_k151_58061	13,603	Caudoviricetes	12	[Prokka] RNA polymerase-associated protein RapA; FMN reductase (NADH) RutF; Protein YrdA; Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex; Dihydrolipoyl dehydrogenase; GDP-perosamine synthase
6_k151_6007	12,135	Caudoviricetes	5	[Prokka] Error-prone DNA polymerase
6_k151_7818	11,554	Caudoviricetes	7	[Prokka] Aerobactin synthase; L-lysine N6-monooxygenase; IS1182 family transposase ISAzch1; Shikimate dehydrogenase (NADP(+))
8_k151_12891	23,115	Caudoviricetes	1	[Prokka] dehydrooxytetracycline reductase; ATP-dependent zinc metalloprotease FtsH; putative sensor histidine kinase TcrY; Transcriptional regulatory protein TcrA; NAD(P)H-quinone oxidoreductase, chloroplastic; NADH-quinone oxidoreductase; Na(+)/H(+) antiporter;
8_k151_26165	12,191	Caudoviricetes	521	[DRAM-v] tRNA-Met(cat); [Prokka] DNA primase; RNA polymerase sigma factor RpoD; Multifunctional non-homologous end joining protein LigD;
8_k151_34356	15,219	Caudoviricetes	4	[Prokka] ATP-dependent RecD-like DNA helicase
8_k151_38164	13,266	Caudoviricetes	1	none identified
8_k151_43	11,891	Caudoviricetes	13	[Prokka] Trehalase; Glucokinase; Transaldolase; Glucose-6-phosphate 1-dehydrogenase 1; 6-phosphogluconate dehydrogenase, NAD(+)-dependent, decarboxylating; Ribulose-phosphate 3-epimerase; Transketolase 2; Ubiquinone biosynthesis O-methyltransferase, mitochondrial
8_k151_60064	23,197	Caudoviricetes	3	[DRAM-v AMG] ABC.PE.P1; peptide/nickel transport system permease protein [Prokka] Glutathione transport system permease protein GsiD; Oligopeptide transport ATP-binding protein OppF; Oligopeptide transport ATP-binding protein OppD; Oligopeptide transport system permease protein OppC; Glutathione transport system permease protein GsiC; Oligopeptide-binding protein OppA; Protease HtpX
8_k151_81666	11,545	Caudoviricetes	58	[Prokka] D-inositol-3-phosphate glycosyltransferase; GDP-mannose 4,6-dehydratase
10_k151_39328	24,398	Caudoviricetes	156	[DRAM-v AMG] GlycosylTransferases [DRAM-v] integrase; putative DNA primase/helicase; adenine-specific DNA-methyltransferase; type III restriction enzyme
10_k151_56258	26,944	Caudoviricetes	8	[DRAM-v] polysaccharide biosynthesis protein PslH; exopolysaccharide production protein ExoZ; adenylate cyclase
10_k151_67091	12,249	Caudoviricetes	32	none identified

ACKNOWLEDGEMENTS

This research is an extension of a project began by Yu Angela Bai and Julie Grace Bevilacque in 2016. Research was supported by the Laidlaw Scholars Foundation, Georgetown University Center for Research and Fellowships, and the Johnson Biosignatures Lab. Research was conducted with the support of Gareth Trubl and Sarah Stewart Johnson. Work conducted at LLNL was conducted under the auspices of the US Department of Energy under Contract DE-AC52-07NA27344.

CONCLUSIONS

- Future research will apply metagenomic analyses to these depth samples with a focus on microbial communities to identify MAGs. Analysis of MAGs may suggest specific virus-host relationships that could allow for observations of the functionality of these interactions in the greater Antarctic soil ecosystem.
- We will continue to focus on the suggested relationship between taxonomic diversity, functional diversity, and depth. In particular, we will further unpack our functional diversity results.
- Data on Antarctic soil microbes will increase understanding of the intricacies of extremophile microbial ecosystems and could have implications for astrobiology research.

References:

1. Yu Angela Bai, Julie Grace Bevilacqua, Aria S. Hahn, Connor Morgan-Lang, Elena Zaikova, Scott Tighe, David Goerlitz, Nicole Wagner, Mia Vanderwilt, Amy McAdam, Christine A. Knudson, Cherie Achilles, Slavka Andrejkovičová, Sarah Stewart Johnson. Selective Pressure of Harsh and Variable Conditions Drives the Decoupling of Taxonomic and Functional Diversity. (unpublished).
2. Evelien M. Adriaenssens, Rolf Kramer, Marc W. Van Goethem, Thulani P. Makhalanyane, Ian Hogg & Don A. Cowan. Environmental drivers of viral community composition in Antarctic soils identified by viromics. *Microbiome* 5, 83 (2017). https://doi.org/10.1186/s40168-017-0301-7
3. Jiarong Guo, Ben Bolduc, Ahmed A. Zayed, Arvind Varsani, Guillermo Dominguez-Huerta, Tom O. Delmont, Akbar Adjie Pratama, M. Consuelo Gazitúa, Dean Vik, Matthew B. Sullivan & Simon Roux. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* 9, 37 (2021). https://doi.org/10.1186/s40168-020-00990-y
4. Antonio Pedro Camargo, Simon Roux, Frederik Schulz, Michal Babinski, Yan Xu, Bin Hu, Patrick S. G. Chain, Stephen Nayfach & Nikos C. Kyrpides. Identification of mobile genetic elements with geNomad. *Nat Biotechnol* (2023). https://doi.org/10.1038/s41587-023-01953-y
5. Michael Shaffer, Mikayla A Borton, Bridget B McGivern, Ahmed A Zayed, Sabina Leanti La Rosa, Lindsey M Solden, Pengfei Liu, Adrienne B Narrowe, Josué Rodríguez-Ramos, Benjamin Bolduc, M Consuelo Gazitúa, Rebecca A Daly, Garrett J Smith, Dean R Vik, Phil B Pope, Matthew B Sullivan, Simon Roux, Kelly C Wrighton. DRAM for distilling microbial metabolism to automate the curation of microbiome function, *Nucleic Acids Research*, Volume 48, Issue 16, 18 September 2020, Pages 8883–8900, https://doi.org/10.1093/nar/gkaa621
6. Ho Bin Jang, Benjamin Bolduc, Olivier Zablocki, Jens H. Kuhn, Simon Roux, Evelien M. Adriaenssens, J. Rodney Brister, Andrew M Kropinski, Mart Krupovic, Rob Lavigne, Dann Turner & Matthew B. Sullivan. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nat Biotechnol* 37, 632–639 (2019). https://doi.org/10.1038/s41587-019-0100-8
7. Roux, S., Camargo, A. P., Coutinho, F. H., Dabdoub, S. M., Dutilh, B. E., Nayfach, S., & Tritt, A. (2023). iPHoP: An integrated machine learning framework to maximize host prediction for metagenome-derived viruses of archaea and bacteria. *PLoS biology*, *21*(4), e3002083. https://doi.org/10.1371/journal.pbio.3002083
8. Torsten Seemann, Prokka: rapid prokaryotic genome annotation, *Bioinformatics*, Volume 30, Issue 14, July 2014, Pages 2068–2069, https://doi.org/10.1093/bioinformatics/btu153
9. Antonio Pedro Camargo, Stephen Nayfach, I-Min A Chen, Krishnaveni Palaniappan, Anna Ratner, Ken Chu, Stephan J Ritter, T B K Reddy, Supratim Mukherjee, Frederik Schulz, Lee Call, Russell Y Neches, Tanja Woyke, Natalia N Ivanova, Emiley A Eloe-Fadrosh, Nikos C Kyrpides, Simon Roux, IMG/VR v4: an expanded database of uncultivated virus genomes within a framework of extensive functional, taxonomic, and ecological metadata, *Nucleic Acids Research*, Volume 51, Issue D1, 6 January 2023, Pages D733–D743, https://doi.org/10.1093/nar/gkac1037