

Atacama Desert Endolithic Communities Harbor Novel and Functional Viruses

Background

Viral metagenomic studies have determined that viruses play important roles in microbial ecology ranging from mediating host metabolism, conferring antibiotic resistance to hosts via horizontal gene transfer, and influencing the structure of communities via infection and phage-induced mortality.

A particularly underexplored area of viral metagenomics is in extreme ecological zones. Between intense solar radiation and the lack of water, deserts are a unique ecological niche inhabited by robust microbial life forms. The Atacama Desert is characterized by its dry and harsh climate with the soil known for its low nutrient content and physical instability (Crits-Christoph, 2016). In contrast, rock substrates termed endoliths retain moisture within their pores and fissures and shield microorganisms within from environmental stressors including desiccation, solar radiation, wind, and temperature fluctuations. As a protective habitat, halite, gypsum, calcite, and ignimbrite rock substrates retain diverse microbial communities.

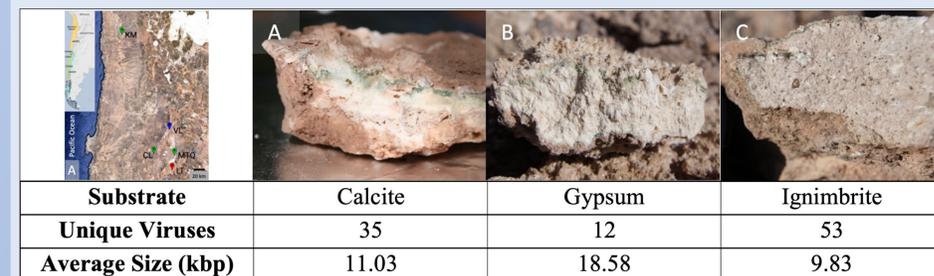


Figure 1: Sampling sites and endoliths with colonization zones from the Atacama Desert (A) Calcite from Valle de la Luna (VL), (B) Gypsum from KM37 (KM), Cordon de Lila (CL), and Monturaqui (MTQ), and (C) Ignimbrite from Lomas de Tilocalar (LT). Viral composition and genome size across the substrates is shown.

While the microbial diversity for endolithic communities has been previously explored, the presence and activity of viruses is unknown. Given the influential capacity viruses exert on microbial communities in other environments, this study evaluated the taxonomic composition, abundance, infective strategies, functional annotations, and phage-host relationships of Atacama Desert endolithic viruses for the first time in the field of microbial ecology.

Methodology

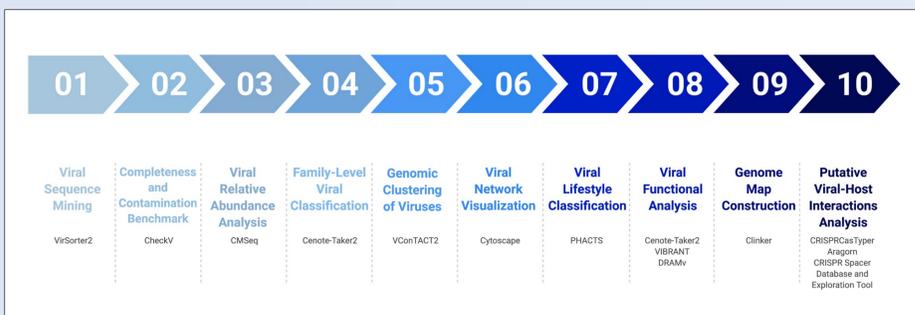


Figure 2: The computational pipeline and tools used to analyze the viromes in this study.

A suite of computational tools were employed to mine the calcite, gypsum, and ignimbrite metagenomes to analyze the viral landscape across the three rock types. VirSorter2 was used to sift through metagenomes from the three endolithic communities for viral contigs. CheckV was then used to remove any contigs identified by VirSorter2 that were shorter than 3 kbp or less than 5% complete. An analysis of the relative abundance, classification, and functional capacity of the viruses was performed with several tools. The computational results in this study set the stage for future experimental verification of phage-host interactions.

Leora Busse and Jocelyne DiRuggiero
Department of Biology

Results

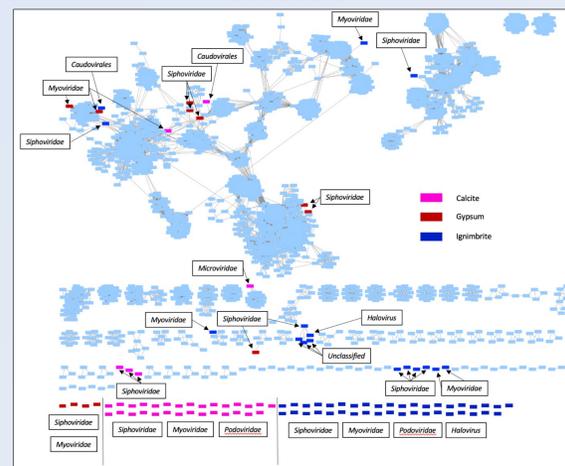


Figure 3: Genome similarity network for calcite, gypsum, ignimbrite, and NCBI RefSeq viruses (light blue). vConTACT2 was used to estimate genus-level taxonomy and Cenote-Taker2 to identify viral families.

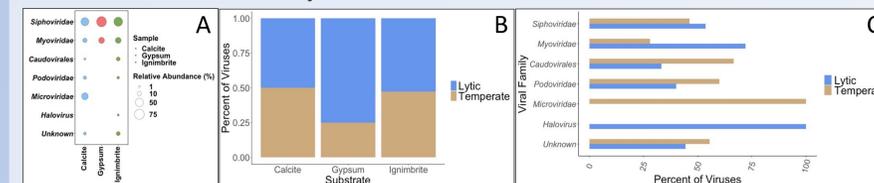


Figure 4: Diversity and life cycles of endolithic viruses. (A) Bubble plot of viral family relative abundance across substrates. (B) Percent of lytic and temperate viruses per substrate. (C) Distribution of viral families with lytic and temperate life cycles in each substrate.

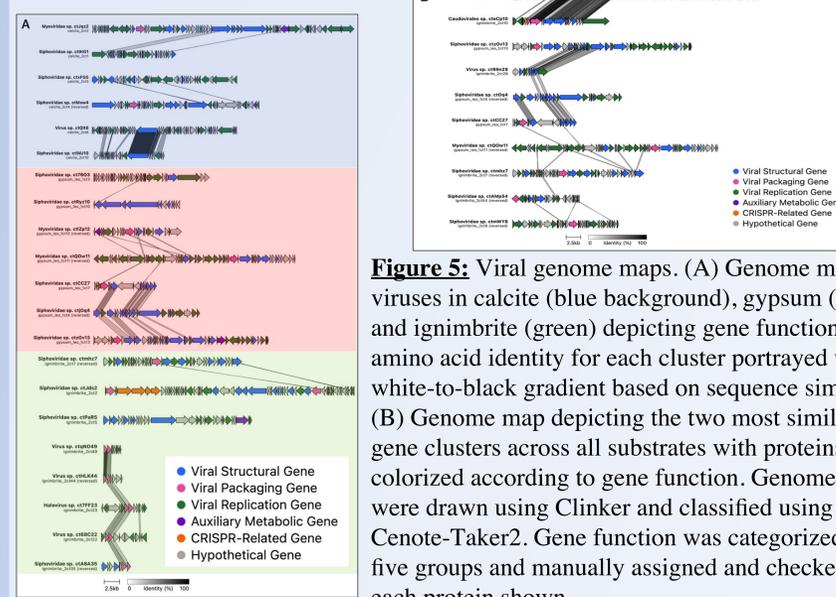


Figure 5: Viral genome maps. (A) Genome maps of viruses in calcite (blue background), gypsum (red), and ignimbrite (green) depicting gene function and amino acid identity for each cluster portrayed with a white-to-black gradient based on sequence similarity. (B) Genome map depicting the two most similar gene clusters across all substrates with proteins colorized according to gene function. Genome maps were drawn using Clinker and classified using Cenote-Taker2. Gene function was categorized into five groups and manually assigned and checked for each protein shown.

Interpretation & Conclusions

Across calcite, gypsum, and ignimbrite microbial communities, *Cyanobacteria* and algae are the primary producers (Meslier, 2018). Of heterotrophic organisms, gypsum and calcite have overall higher levels of microbial diversity compared to ignimbrite.

This study identified 100 unique viruses in the three endoliths (Fig. 1). Only 30 formed clusters with each other or to viruses in the reference database (Fig. 3). The other 70 viruses share no genomic similarity to other viruses and represent novel viral sequences. Diversity metrics suggest that the ignimbrite and calcite metaviromes are highly diverse compared to the gypsum (Fig. 4A). Overall, *Siphoviridae* dominates the three endoliths while other families are present in lower relative abundance.

Assessing viral infective capabilities across the endoliths revealed no preference for either the lytic (active infection in host cells) or temperate (DNA integration in host cells) life cycle in the calcite or ignimbrite while there is a notable increase in lytic viruses in the gypsum (Fig. 4B). Viruses in warm deserts tend to be temperate due to harsh environmental factors; however, more lytic viruses have been identified in Atacama soil and halite viromes (Hwang, 2021, Crits-Christoph, 2013). Protection in endoliths from environmental stressors may enhance viral active infection. Within certain viral families, there is a noted preference for one cycle over the other (Fig. 4B). Viral genetics and host immunity may influence the life cycle prevalence in each viral family.

The lack of sequence similarity in the viral metagenome-assembled genomes (vMAGs) suggests that viruses within and across the endoliths are structurally novel. Most proteins share little to no amino acid identity with other viruses; however, viral structural and replication proteins in the calcite and gypsum and DNA packaging proteins in the ignimbrite tend to share the highest similarity overall (Fig. 5).

Viruses hijack auxiliary metabolic genes (AMGs) from bacteria to augment viral replication by mediating host metabolism. 20 AMGs were found across 11 endolithic viruses, with carbohydrate, vitamin, and amino acid metabolism spanning all three substrates. Substrate-specific AMGs retain other functions (Fig. 6).

Lastly, phage-host relationships are predicted via genomic similarity of CRISPR spacer and tRNA sequences. Across the vMAGs, 2 ignimbrite viruses contained CRISPR proteins or spacer arrays, and 6 viruses overall contained tRNAs.

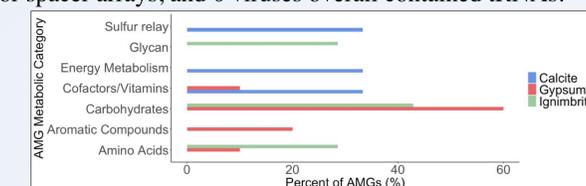


Figure 6: Categorized auxiliary metabolic genes for each rock substrate as identified by VIBRANT.

Overall, 100 novel and functional viruses were identified in Atacama Desert endoliths. Viruses retain important roles in shaping community dynamics, and the presence of AMGs and spacer arrays suggests viral infection occurs in the studied microbial communities. This project enhances the understanding of endolithic viral activity in the field of microbial ecology.

Acknowledgements

Special thanks to Dr. Jocelyne DiRuggiero for her dedicated mentorship on this project, the URSCA Office, Dr. Mike Tisza, Dr. Natalie Strobach, Dr. Julie Lirot, and Dr. Caitlin Larracey for their guidance on my research and future career.

This work was funded by the Woodrow Wilson Undergraduate Research Fellowship to LB and by NSF grant DEB1556574 and NASA grant NNX15AP18G to JDR.

Crits-Christoph, A. 2016. "Functional Interactions of Archaea, Bacteria and Viruses in a Hypersaline Endolithic Community." *Env. Microbiol.*, doi: 10.1111/1462-1462-2920.13259.

Ertekin E. 2020. Functional and taxonomic diversity is driven by substrate architecture in endolithic communities from extreme environments. *Env. Microbiol.* <https://doi.org/10.1111/1462-2920.15287>.

Meslier V. 2018. Fundamental drivers for endolithic microbial community assemblies in the hyperarid Atacama Desert. *Env Microbiol*20:1765-1781 doi: 10.1111/1462-2920.14106.