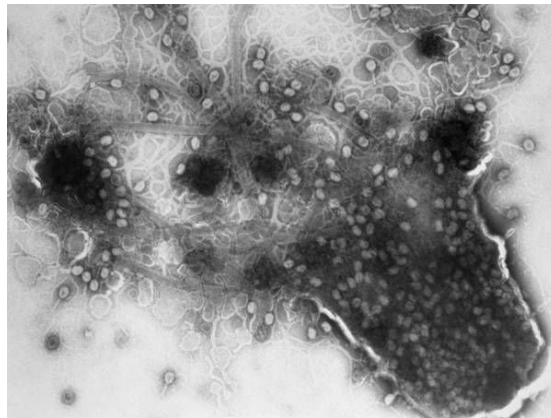


# Characterizing prophages in prokaryotic genomes from a various environments and across the tree of life

Project 2020



A bacterial cell being lysed from phage infection. Image source:(<https://fineartamerica.com/>)

Contact:

Migun Shakya, PhD

[migun@lanl.gov](mailto:migun@lanl.gov)

Brett Youtsey

[byoutsey@lanl.gov](mailto:byoutsey@lanl.gov)

## Background:

Prophages, sometimes also referred to as proviruses are viruses that integrate themselves into a host bacterial and archaeal genomes. These prophages become part of their host genome and stay integrated until the conditions become unfavourable such as host cell damage, lack of nutrients, etc. Such conditions trigger a lytic life cycle where prophages hijack their host and use their resources to replicate their genome, make more virus particles, only pack their genome, and then lyse and kill the cells to escape and to look for another host. This fascinating life cycle has many consequences that range from affecting virulence of pathogens, playing an important role in Horizontal Gene Transfer, and contributing to the global carbon cycle.

In this project, we will characterize prophages from publicly available bacterial and archaeal genomes that have been isolated and sequenced from various environments ranging from soil to marine. We are specifically interested in understanding if bacteria residing in vastly different environments have significantly different numbers of prophages. As a part of the

project, each member of the group will be assigned to a specific environment or a taxonomic group and individuals will coordinate with each other to work on the assigned environment and will convene at the end to compare and interpret their results and further discuss differences and similarities of prophage distribution.

## What you'd be working on:

- Mining publicly available databases for selecting and downloading thousands of genomes. Specifically, using Entrez Utilities to select bacterial genomes from GenBank and RefSeq.
- Running third party phage search tools (PhiSpy, PHASTEST, VirSorter, etc) on bacterial genomes.
- Genome analyses using Python or R to track the position of prophages, calculate GC content, visualize the results etc.
- Statistical tests to infer differences in frequency of prophages among different environments and taxonomic groups.
- Constructing phylogeny of host genomes.

## Suggested Readings:

1. Paez-Espino, D., Eloie-Fadrosch, E. A., Pavlopoulos, G. A., Thomas, A. D., Huntemann, M., Mikhailova, N., ... & Kyrpides, N. C. (2016). Uncovering Earth's virome. *Nature*, 536(7617), 425-430.
2. Touchon, M., Bernheim, A., & Rocha, E. P. (2016). Genetic and life-history traits associated with the distribution of prophages in bacteria. *The ISME journal*, 10(11), 2744-2754.
3. Life in our phage world  
book:<https://drive.google.com/file/d/0BxYktwFjAkFpUkczODFH1BPSHc/view>