**SESSION GUIDE:
Collaborative Annotation Teams (CATs): CLUSTER REPORTS**

**PURPOSE**:
To capture our current understanding of the characteristics of a given cluster, that can be used to support the process of genome annotations, genome announcement preparations, inform and contribute towards future studies examining phages of this cluster.

**INSTRUCTIONS:**

* Use the sample/template Cluster Report below to develop a Cluster Report for your cluster.
*Note: Text in green is information specific to the cluster is the report. To use this template, replace the text in green with information relevant to the cluster you are examining. The text in black providing framing for the types of information that should be provided, and can be edited as necessary.*
* To facilitate your ability to gather the relevant information for your cluster, the Cresawn lab has developed an Observable Notebook that draws data from Phamerator and phagesDB.
Link to Observable Notebook: <https://observablehq.com/d/5e5bc78c9b3ae2ed>
* Work collaboratively to develop a single report, guided by the SMART members in your group.
* At the end of the session,
	+ The SMART/Expedited Submitter should upload the document to the CAT Forum:
	Access the CAT forums here: <https://seaphages.org/forums/forum/228/>

Over the coming months, these Cluster Reports will be reviewed and then published to QUBES, recognizing you, the authors, with a citable publications and available to advance the work of this community.

**A 2024 Report on the Characteristics of Phages Assigned to Cluster D1.**

**Authors**:

**Author Affiliations**:

**Reviewing Author**: (typically, the member of the SMART Team assigned to this cluster)
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Isolation and Morphological Characteristics**

As of June 2024, 21 phages have been assigned to Actinobacteriophage cluster D1. Cluster D1 consists of lytic phages that were isolated from environmental samples collected at 8 different states, using *Mycobacterium smegmatis* as the host. A sampling of transmission electron micrographs reveal these phages to have siphoviral morphologies, while photographs of plaque assays reveal small- to medium-sized clear plaques. Examples of phages displaying these morphologies include phage Chill and Gumball.

**Genometrics**
The genomes of phages in this cluster have an average length of 64,676 bp (range: 64,494 bp – 65,108 bp) with an average GC content of 59.7% (range: 59.6% – 59.8%). The genomes have circularly permuted ends.

**Gene Content**

1. A survey of all manually annotated genomes in this cluster (i.e. excluding draft genomes) reveal the number of genes called per genome to range between 81 – 91 genes, which represent a total of 115 gene phamilies (or phams).
2. Of these 115 phams, 66 are fully conserved across the cluster. These phams are assigned putative functions that are involved in structure and assembly (terminase, portal protein, major and minor capsid proteins, tape measure, tail assembly chaperones, and minor tail proteins), replication and recombination (DNA primase/polymerase, DNA polymerase alpha subunit, DNA helicase, and resolvase), and lysis (lysinA and lysinB). Some other notable putative functions include ThyX-like thymidylate synthase, FabG-like reductase, metal transferase, and oxidoreductase.
3. Another 48 phams have varying levels of conservation. Notable here are 9 phams that are each only called once in this cluster, while another 6 phams are each only called twice in this cluster. One of these latter phams is assigned as a VIP2-like toxin domain.
4. There are no tRNA genes identified in this cluster. There are also no identifiable integrase or repressor functions, suggesting that phages of this cluster do not form lysogens and are thus lytic phages.
5. The genomes are organized with genes involved in structure and assembly located within the first third of the genome while those involved in replication and recombination are scattered across the remaining two-thirds of the genome. Genes involved in lysis are located at the boundary of these two regions. All genes are transcribed in the forward direction, with the exception of 3 highly conserved genes adjacent to the lysis gene that are transcribed in the reserve direction. In phage BigMama, two genes have been called that break this pattern of gene arrangement across this cluster.

**References:** Below are prior publications that include a description of this cluster (consider excluding genome announcements) **Acknowledgements:** We thank Maddie Bendele, Isobel Cobb, and Steve Cresawn for developing the tools that were used to generate this report.