A 2024 Report on the Characteristics of Phages Assigned to Cluster CS2

Authors: Ping An[1], Dane Bowder[2]

Reviewing Author: Marcie Warner[1]

Author Affiliations: [1]University of Pittsburgh; [2]Doane University

---

Isolation and Morphological Characteristics

As of October 2024, 31 annotated phages have been assigned to Actinobacteriophage cluster CS2. Cluster CS2 consists of lytic phages isolated from environmental samples collected in the United States of America from 11 different states, using *Gordonia terrae* and *Gordonia rubripertincta* as the host. A sampling of transmission electron micrographs reveals these phages to have siphoviral morphologies, while photographs of plaque assays reveal small- to medium-sized clear plaques, some phages appear to form multiple plaque sizes on the same plate. Examples of phages displaying these morphologies include phages CathyBurgh, Chelms, and BirksandSocks.

Genometrics

The genomes of phages in this cluster have an average length of 76,603 bp (range: 74934 bp – 78164 bp) with an average GC content of 58.9 % (range: 58.6% – 59.3%). The genomes of 97% of CS2 phages have direct terminal repeat ends. No specific end type was identified for 3% of CS2 phages.

Gene Content

a. A survey of all manually annotated genomes in this cluster (i.e., excluding draft genomes) reveals the number of genes called per genome ranges between 103-110 genes, which represent a total of 161 gene phamilies (or phams).

b. Of these 161 phams, 75 are fully conserved across the cluster. These phams are assigned putative functions that are involved in structure and assembly (e.g., terminase, portal protein, major and minor capsid proteins, tape measure, tail assembly chaperones, head-to-tail adaptor, capsid maturation protease, capsid decoration protein and minor tail proteins), replication and recombination (e.g., DNA primase/polymerase, DnaQ-like DNA Polymerase III subunit, DNA polymerase III sliding clamp, DNA helicase, and

resolvase), and lysis (e.g., lysin A, N-acetylmuramoyl-L-alanine amidase domain, M23 domain, lysin B and holin). Some other notable putative functions include CobT-like cobalamin biosynthesis protein, DNA methylase, Cas4 family exonuclease, AAA-ATPase, RusA-like resolvase, and nucleoside deoxyribosyltransferase.

c. Another 86 phams have varying levels of conservation. Notable here are 18 phams that are each only called once in this cluster across 10 phages. Among these 18 phams, Pham 191237 is called in SteveFrench and is associated with the function HNH endonuclease. Another 5 phams are each only called twice in this cluster. Among these 5 phams, Pham 12647 is called in RoyalG and SteveFrench, and has an esterase function.
d. There is one tRNA gene identified in this cluster. All manually annotated CS2 phages encode a tRNA-Asn; presently, 1 phage (PokyPuppy) has a tRNA-Gln, but a review of the data suggests this should be corrected to Asn – a correction has been submitted). There are no identifiable integrase or repressor functions, suggesting that phages of this cluster do not form lysogens and are, thus, lytic phages.
e. The genomes are organized with genes involved in structure and assembly located within the first half of the genome, while those involved in replication and recombination are scattered across the remaining half of the genome. Genes involved in lysis are located at the boundary of these two regions. There is a clear forward and reverse arm in these genomes with a break about halfway across the genome, with the exception of 8 reverse genes found at the 5’ end of the genome upstream of all the forward genes.

References

1. Pope WH, Mavrich TN, Garlena RA, Guerrero-Bustamante CA, Jacobs-Sera D, Montgomery MT, Russell DA, Warner MH, Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES), Hatfull GF. 2017. Bacteriophages of Gordonia spp. Display a Spectrum of Diversity and Genetic Relationships. mBio 8:10.1128/mbio.01069-17. https://doi.org/10.1128/mbio.01069-17

Source Data: https://observablehq.com/d/5e5bc78c9b3ae2ed accessed July 2024.

Acknowledgments: We thank Maddie Bendele, Isobel Cobb, and Steve Cresawn for developing the tools used to generate this report.