A 2024 Report on the Characteristics of Phages Assigned to Subcluster AU6

Authors: Amaya Garcia Costas[1] and Quinn Vega[2]

Reviewing Author: Karen Klyczek[3]

Author Affiliations: [1]Colorado State University Pueblo, Pueblo, CO; [2]Montclair State University, Montclair, NJ; [3]University of Wisconsin-River Falls (emeritus), River Falls, WI

Isolation and Morphological Characteristics

As of July 2024, 8 phages have been assigned to Actinobacteriophage subcluster AU6 (5 are annotated, 3 are drafts). Subcluster AU6 consists of lytic phages isolated from environmental samples collected in 3 US states (NY, MA, CO) and 2 Canadian provinces (Ontario, Quebec), using *Arthrobacter globiformis* as the host. A sampling of transmission electron micrographs reveals that 25% of phages have siphoviral morphologies; the remaining 75% of phages in the subcluster do not have images on phagesdb.org. Photographs of plaque assays reveal medium plaques that are completely clear or have a clear center surrounded by a halo. Examples of phages displaying these morphologies include GantcherGoblin and Zeina.

Genometrics

The genomes of phages in this subcluster have an average length of 55,238 bp (range: 55,017 bp – 55,643 bp) with an average GC content of 50.5% (range: 49.7% – 51.2%). The genomes have 3’ sticky ends with 9 base overhangs.

Gene Content

a. A survey of the five manually annotated genomes in this subcluster (i.e., excluding draft genomes) reveals that the number of genes called per genome ranges between 89-92 genes, with 91.2 genes per genome on average. There are 438 total genes in these annotated genomes of subcluster AU6, which represent a total of 120 gene phamilies (or phams).

b. Of the 120 phams represented in the subcluster, 63 are fully conserved. These phams are assigned putative functions that are involved in structure and assembly (e.g., terminase, portal protein, major capsid protein, tape measure, tail terminator, major tail protein, and various minor tail proteins), replication and recombination (e.g., DNA primase/polymerase, SSB protein/SSB DNA binding protein, DNA helicase), and lysis (e.g., holin, endolysin/lysinA). Some other notable putative functions include helix-turn-helix DNA binding domain protein, several HNH endonucleases, Cas4 family exonuclease, and a PnuC-like nicotinamide riboside transporter.

c. The other 57 phams have varying levels of conservation. Of these, 22 phams are each only called once in this subcluster, while another 14 phams are each only called twice in the subcluster. These all have hypothetical function assigned to them. Two phams, with

acetyltransferase and minor tail functions, are called in all but one of the five annotated genomes. Two phams are not conserved among all annotated members of the subcluster but are assigned the same function, VRR Nuc domain protein, making the function conserved across the annotated genomes.

d. No tRNA genes have been identified in this subcluster. There are also no identifiable integrase or repressor functions, suggesting that phages of this subcluster likely 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. Within the structural genes, there is also an acetyltransferase and endolysin/lysinA, typically found in this subcluster at the very left of the genome. There are genes with hypothetical assigned function distributed across the genome.

References: No prior publications**.**

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.