

# Complete Genome Sequence of the Cluster DJ Actinobacteriophage, *Petito*, isolated on the host *Gordonia rubripertincta*

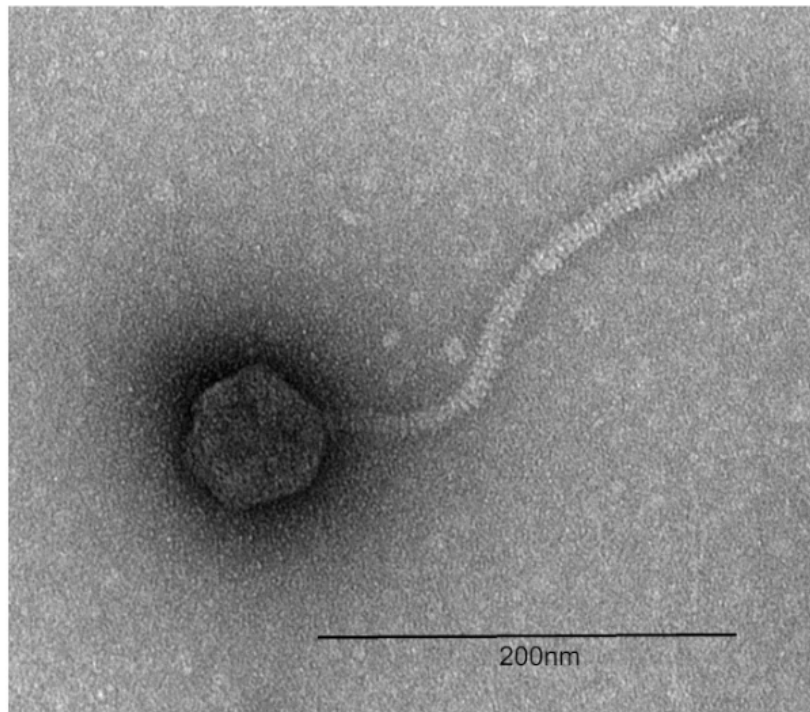
Elliot Benson<sup>1</sup>, Micah Blount<sup>1</sup>, Shria Chauhan<sup>1</sup>, Jayden Ehrhart<sup>1</sup>, Adelinn Foster<sup>1</sup>, Abigail Ingber<sup>1</sup>, Madeline Julian<sup>1</sup>, Derika Kwansah<sup>1</sup>, Trang Le<sup>1</sup>, Emily May<sup>1</sup>, Elizabeth Mazel<sup>1</sup>, Esther Morency<sup>1</sup>, Sierra Nelson<sup>1</sup>, Casey O'Toole<sup>1</sup>, Kaitlin Potter<sup>1</sup>, Leandra Vita<sup>1</sup>, Kirra Weigand<sup>1</sup>, Denise Monti<sup>1§</sup>

<sup>1</sup>Hicks Honors College, University of North Florida, Jacksonville, Florida, United States

<sup>§</sup>To whom correspondence should be addressed: montid@hhmi.org

## Abstract

*Gordonia* phage *Petito* is a newly discovered siphovirus that infects *Gordonia rubripertincta* NRRL B-16540. The double-stranded DNA genome of this phage is 60,447 bp long with 93 predicted protein-coding genes and no tRNAs. *Petito* is a Cluster DJ phage.



**Figure 1. Transmission electron micrograph of *Gordonia* phage *Petito*.**

Phage lysate was negatively stained with 1% uranyl acetate and imaged with a Hitachi HT7800 120kV TEM and an AMT Nanospring15 B digital camera. Scale bar = 200nm.

## Description

Bacteria that comprise the genus *Gordonia* are metabolically diverse and have been implicated as possible agents for bioremediation (Arenskötter et al. 2004). *Gordonia* spp. may have additional use in biotechnology as biomachines for synthesizing novel synthetic compounds. To better understand *Gordonia* spp. and to identify novel genes that may be exploited for synthetic biology, students in the Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program have isolated and characterized nearly 800 phages from nine species of *Gordonia* (phagesdb.org) (Russell 2016). Here, we report the isolation and characterization of phage *Petito* isolated on the host, *Gordonia rubripertincta*.

*Petito* was isolated from soil collected from Ogier Gardens at the University of North Florida, Jacksonville, FL (30.27945 N, 81.51295 W). Soil samples were washed with Luria Broth (LB), incubated at room temperature with shaking for 1 h, and

filtered through a 0.22 $\mu$ M filter. Soil extracts were enriched for phages infecting *G. rubripertincta* by inoculating the extracts with *G. rubripertincta* NRRL B-16540 and incubating for 48 hours at 30°C. Extracts were then filtered through a 0.22 $\mu$ M filter, 10-fold serially diluted, and plated with LB top agar containing *G. rubripertincta* NRRL B-16540. Phage replication formed small, clear plaques when incubated at 30°C for 24 hours. Petito was purified by three rounds of plaque purification assay (Zorawik 2024), after which a viral lysate was prepared. Negatively stained transmission electron microscopy revealed siphoviral morphology with an icosahedral capsid ( $69 \pm 2$  nm diameter,  $n=2$ ) and a noncontractile tail ( $252 \pm 0.8$  nm length,  $n=2$ ).

Phage DNA was extracted using the Wizard DNA cleanup kit (Promega). A sequencing library was prepared using the NEB Ultra II Library Kit and sequenced using an Illumina MiSeq (v3 Reagents), yielding 298,983 single-end 150-based single-end reads (~719-fold coverage). Raw reads assembled and checked using Newbler v2.9 (Russell 2017) and Consed v29 (Gordon and Green 2013) with the default parameters yielded a single contig 60,447bp in length with 3' single-stranded overhangs (5' CGCCGTCCT).

DNAMaster (v5.23.6) and Phage Evidence Collection and Annotation Network (PECAAN) (<https://blog.kbrinsgd.org/overview/>) were used to annotate the genome. Ninety-three protein-coding genes were identified based on predictions by Glimmer v3.02 (Delcher et al. 2007), GeneMark v2.0 (Besemer and Borodovsky 2005), Starterator (v.575), and BLASTp against the Actinobacteriophage protein database. The genome was scanned for tRNAs using ARAGORN v1.2.38 (Laslett and Canback 2004). No tRNAs were identified. Protein functions were assigned using HHPred (PDB\_mmCIF70, SCOPe70, Pfam-A v36, NCBI Conserved Domain database) (Söding 2005), BLASTp v2.14.1 alignments against the NCBI non-redundant protein sequences databases (<https://blast.ncbi.nlm.nih.gov>), and Phamerator (v575) (Cresawn et al., 2011). Transmembrane domains were identified by DeepTMHMM v1.0.24 (<https://dtu.biolibn.com/DeepTMHMM>) (Hallgren et al. 2022). Putative protein-coding genes were sorted into Phams using the pipeline PhaMMseqs (Gauthier, 2022). All software was used with default parameters. Petito was assigned to Cluster DJ based on at least 35% gene content similarity (GCS) to sequenced phages in the Actinobacteriophage database (Gauthier and Hatfull, 2023).

Similar to other phages assigned to Cluster DJ, all genes in Petito are transcribed in the forward direction. Twenty-four genes were assigned a putative function. Genes for structure, assembly, and lysis are encoded in the first third of the genome. Two copies of the major tail protein are encoded in the left arm of the genome along with other structure-related genes (e.g., minor tail protein, tail terminator). Of interest are two homologs of unknown function (Petito gp50 and gp61) encoded in the center of the genome that are also found in Arthrobacter, Microbacterium, and Streptomyces phages. A second pair of homologs of unknown function (Petito gp81 and gp82) are encoded in the last third of the genome. Unlike Petito gp50 and gp61, Petito gp81 and gp82 are specific to Cluster DJ phages.

#### Nucleotide sequence accession numbers

Petito is available at GenBank with Accession No. [PP758912](https://www.ncbi.nlm.nih.gov/nuclseq/PP758912) and Sequence Read Archive (SRA) No. 30907824.

**Acknowledgements:** We thank Anjali O'Toole Gupta for assistance in phage isolation, the Pittsburgh Bacteriophage Institute and the HHMI Science Education Alliance for phage sequencing. We also thank Tagide deCarvalho at the University of Maryland Baltimore County Keith R. Porter Imaging Facility for the electron microscopy images.

#### References

- Arenskötter M, Bröker D, Steinbüchel A. 2004. Biology of the metabolically diverse genus *Gordonia*. *Appl Environ Microbiol* 70(6): 3195-204. PubMed ID: [15184112](https://pubmed.ncbi.nlm.nih.gov/15184112/)
- Besemer J, Borodovsky M. 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res* 33(Web Server issue): W451-4. PubMed ID: [15980510](https://pubmed.ncbi.nlm.nih.gov/15980510/)
- Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF. 2011. Phamerator: a bioinformatic tool for comparative bacteriophage genomics. *BMC Bioinformatics* 12: 395. PubMed ID: [21991981](https://pubmed.ncbi.nlm.nih.gov/21991981/)
- Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23(6): 673-9. PubMed ID: [17237039](https://pubmed.ncbi.nlm.nih.gov/17237039/)
- Gauthier CH, Cresawn SG, Hatfull GF. 2022. PhaMMseqs: a new pipeline for constructing phage gene families using MMseqs2. *G3 (Bethesda)* 12(11). PubMed ID: [36161315](https://pubmed.ncbi.nlm.nih.gov/36161315/)
- Gauthier CH, Hatfull GF. 2023. PhamClust: a phage genome clustering tool using proteomic equivalence. *mSystems* 8(5): e0044323. PubMed ID: [37791778](https://pubmed.ncbi.nlm.nih.gov/37791778/)

Gordon D, Green P. 2013. Consed: a graphical editor for next-generation sequencing. *Bioinformatics* 29(22): 2936-7. PubMed ID: [23995391](#)

Hallgren J, Tsirigos KD, Pedersen MD, Almagro Armenteros JJ, Marcatili P, Nielsen H, Krogh A, Winther O. 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. : 10.1101/2022.04.08.487609. DOI: [10.1101/2022.04.08.487609](#)

Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 32(1): 11-6. PubMed ID: [14704338](#)

Russell DA. 2017. Sequencing, Assembling, and Finishing Complete Bacteriophage Genomes. *Methods in Molecular Biology, Bacteriophages* : 109-125. DOI: [10.1007/978-1-4939-7343-9\\_9](#)

Russell DA, Hatfull GF. 2016. PhagesDB: the actinobacteriophage database. *Bioinformatics* 33: 784-786. DOI: [10.1093/bioinformatics/btw711](#)

Soding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research* 33: W244-W248. DOI: [10.1093/nar/gki408](#)

Zorawik M, Jacobs-Sera D, Freise AC, SEA-PHAGES, Reddi K. 2024. Isolation of Bacteriophages on Actinobacteria Hosts. *Methods Mol Biol* 2793: 273-298. PubMed ID: [38526736](#)

**Funding:** This work was supported by a UNF Foundation Faculty Grant to D. Monti.

**Author Contributions:** Elliot Benson: investigation, writing - original draft. Micah Blount: investigation, writing - original draft. Shria Chauhan: investigation, writing - original draft. Jayden Ehrhart: investigation, writing - original draft. Adelinn Foster: investigation, writing - original draft. Abigail Ingber: investigation, writing - original draft. Madeline Julian: investigation, writing - original draft. Derika Kwansah: investigation, writing - original draft. Trang Le: investigation, writing - original draft. Emily May: investigation, writing - original draft. Elizabeth Mazel: investigation, writing - original draft. Esther Morency: investigation, writing - original draft. Sierra Nelson: investigation, writing - original draft. Casey O'Toole: investigation, writing - original draft. Kaitlin Potter: investigation, writing - original draft. Leandra Vita: investigation, writing - original draft. Kirra Weigand: investigation, writing - original draft. Denise Monti: funding acquisition, supervision, visualization, writing - review editing.

**Reviewed By:** Anonymous

**History:** Received December 11, 2024 **Revision Received** January 3, 2025 **Accepted** January 3, 2025 **Published Online** January 3, 2025 **Indexed** January 17, 2025

**Copyright:** © 2025 by the authors. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Citation:** Benson, E; Blount, M; Chauhan, S; Ehrhart, J; Foster, A; Ingber, A; et al.; Monti, D (2025). Complete Genome Sequence of the Cluster DJ Actinobacteriophage, *Petito*, isolated on the host *Gordonia rubripertinca*. *microPublication Biology*. [10.17912/micropub.biology.001456](#)