

Complete Genome Sequences of seven cluster DE1 phages isolated using *Gordonia terrae*

Azzam Almutairi¹, Gabriella Marsala², Natalie Mershon³, Sarah Ball^{4§}

¹Department of Biochemistry, The Ohio State University, Columbus, Ohio, United States

²Department of Psychology, The Ohio State University, Columbus, Ohio, United States

³Center for Life Sciences Education, The Ohio State University, Columbus, Ohio, United States

⁴Department of Molecular Genetics, The Ohio State University, Columbus, Ohio, United States

[§]To whom correspondence should be addressed: ball.1766@osu.edu

Abstract

The phages Barsten, Kamashten, Kwobi, LilHam, SchottB, StorminNorm and Thing3 were isolated from soil samples collected around Columbus, Ohio, United States, using the host *Gordonia terrae* 3612. The seven genomes range in size between 55,593bp and 59,501bp, and contain an average of 85 protein-coding genes. Based on gene content similarity, all seven phages are assigned to actinobacteriophage cluster DE1.

Phage	GPS coordinates	Approximate read coverage	# of 150 base single-end reads	Genome length (bp)	GC content %	# of ORFs (% with predicted function)
Barsten	41.477357 N, 81.626524 W	3261	1,466,780	59501	67.70%	88 (26%)
Kamashten	40.003044 N, 83.027215 W	559	231,687	59294	67.40%	86 (28%)
Kwobi	39.997355 N, 83.014485 W	842	359,399	58482	68.10%	87 (34%)
LilHam	39.995022 N, 83.015197 W	756	371,760	55593	68.20%	80 (29%)
SchottB	40.006111 N, 83.021389 W	118	45,597	58367	68.00%	85 (28%)
StorminNorm	39.88 N, 82.74 W	100	38,931	58248	68.30%	86 (30%)
Thing3	40.001094 N, 83.004358 W	904	379,397	59294	67.40%	86 (35%)

Figure 1. Genomic characteristics of isolated phages:

The sequencing and genome information for seven DE1 phages isolated using *Gordonia terrae* 3612, including read coverage and number of 150 base single-end reads, genome length, GC content, and number of predicted open reading frames.

Description

Gordonia spp. are gram-positive bacteria belonging to the phylum Actinobacteria. Some members of this genus may have environmental and biotechnical uses due to their ability to break down natural polymers and pollutants, while others are opportunistic pathogens (Arenskötter et al. 2004). To better understand the diversity of bacteriophages capable of infecting *Gordonia spp.*, seven phages were isolated and purified from soil samples using the host bacterium *Gordonia terrae* 3612.

All seven of the phages were isolated from soil samples collected in Columbus, Ohio, United States, following standard protocols (Zorawik et al. 2024)). Isolation was achieved by shaking the soil sample with peptone-yeast extract-calcium (PYCa) liquid media for 1 hour, then centrifuging the samples for 10 minutes at 2,000 x g. The supernatant was filtered through a 0.22 µm filter and plated in PYCa top agar with *Gordonia terrae* 3612. The plates were incubated at 30°C for 7 days. The phages were purified through three rounds of plaque streaking. Following incubation, StorminNorm and Thing3 produced slightly turbid plaques, while the other phages produced clear plaques. Plaque sizes ranged from 2.0 to 3.0mm. All the phages exhibit siphovirus morphology.

Phage DNA was extracted from a lysate using a Promega Wizard DNA Clean-Up kit. The DNA was prepared for sequencing with the NEB Ultra II Library Kit and sequenced with an Illumina MiSeq 1000 sequencer (v3 reagents). The raw reads were assembled by Newbler v2.9 (Miller et al., 2010) with default settings and Consed v2.9 (Gordon and Green, 2013) was used to check accuracy, coverage, and genomic termini (Figure 1). The result was genomes with circularly permuted ends. All seven phages were assigned to cluster DE1 based on gene content similarity (GCS) above 35% to phages within the Actinobacteriophage Database (Russell & Hatfull, 2017; Pope et al. 2017). The genomes ranged from 59,501 bp to 55,593 bp in length, with an average GC content of 67.9% (Figure 1).

The genomes were automatically annotated in DNAMaster v5.23.6 (Pope and Jacobs-Sera, 2018) using Glimmer v3.0 (Delcher et al. 2007) and GeneMark v2.5 (Borodovsky, 2005). The annotations were refined through BLAST searches against the NCBI nonredundant and actinobacteriophage databases (Altschul et al., 1990), and HHpred searches against these four databases: PDB_mmCIF70, SCOPe70, Pfam-A, and NCBI_Conserved_Domains (CD) (Söding et al., 2005). Phamerator (Cresawn et al., 2011), using the Actino_draft database and TMHMM were also employed, with default settings. No tRNAs were detected with Aragorn v1.2.38 (Laslett and Canback, 2004) or tRNAscan-SE v2.0 (Lowe and Eddy, 1997). The percentage of genes assigned a putative function ranged from 26%-35% (Figure 1). The absence of an integrase gene supports the prediction that all seven phages are lytic and cannot produce lysogens, a feature shared by members of the DE1 cluster. LilHam and SchottB each contained one orpham (gene with no known homologs), while the other phages contained none. All predicted protein-coding genes are transcribed in the same direction, with no significant gaps between open reading frames. All seven genomes display the typical order of genes found in cluster DE1, with structural genes on the left arm, followed by various tail proteins and the lysis cassette, which includes both lysin A and lysin B. The nucleotide similarity of the right arm of the genomes shows less conservation between the phages, however they all contain genes involved in DNA metabolism.

Nucleotide sequence accession numbers

Barsten is available at GenBank Accession No. MT498035 and Sequence Read Archive (SRA) No. MT498035

Kamashten is available at GenBank Accession No. PV915807 and Sequence Read Archive (SRA) No. PV915807

Kwobi is available at GenBank Accession No. PV915880 and Sequence Read Archive (SRA) No. PV915880

LilHam is available at GenBank Accession No. PV915824 and Sequence Read Archive (SRA) No. PV915824

SchottB is available at GenBank Accession No. OQ995426 and Sequence Read Archive (SRA) No. OQ995426

StorminNorm is available at GenBank Accession No. OQ995426 and Sequence Read Archive (SRA) No. OQ995426

Thing3 is available at GenBank Accession No. OQ995426 and Sequence Read Archive (SRA) No. OQ995426

Acknowledgements: We'd like to thank the Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program for sequencing the genomes. We also thank the other SEA-PHAGES students at The Ohio State University who participated in the annotation of these genomes, as well as Dr. Chuck Daniels and Dr. Greg Booton who helped to refine the annotations and proofread the manuscript.

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403-410. DOI: [10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- 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 Research* 33: W451-W454. DOI: [10.1093/nar/gki487](https://doi.org/10.1093/nar/gki487)
- 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: 10.1186/1471-2105-12-395. DOI: [10.1186/1471-2105-12-395](https://doi.org/10.1186/1471-2105-12-395)
- Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23: 673-679. DOI: [10.1093/bioinformatics/btm009](https://doi.org/10.1093/bioinformatics/btm009)
- Gordon D, Green P. 2013. *Consed*: a graphical editor for next-generation sequencing. *Bioinformatics* 29: 2936-2937. DOI: [10.1093/bioinformatics/btt515](https://doi.org/10.1093/bioinformatics/btt515)
- Hallgren J, Tsirigou 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. *bioRxiv* 10.1101/2022.04.08.487609. DOI: [10.1101/2022.04.08.487609](https://doi.org/10.1101/2022.04.08.487609)

Laslett D. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Research 32: 11-16. DOI: [10.1093/nar/gkh152](https://doi.org/10.1093/nar/gkh152)

Lowe TM, Eddy SR. 1997. tRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence. Nucleic Acids Research 25: 955-964. DOI: [10.1093/nar/25.5.0955](https://doi.org/10.1093/nar/25.5.0955)

Miller JR, Koren S, Sutton G. 2010. Assembly algorithms for next-generation sequencing data. Genomics 95: 315-327. DOI: [10.1016/j.ygeno.2010.03.001](https://doi.org/10.1016/j.ygeno.2010.03.001)

Pope WH, Jacobs-Sera D. 2018. Annotation of Bacteriophage Genome Sequences Using DNA Master: An Overview. Methods Mol Biol 1681: 217-229. PubMed ID: [29134598](https://pubmed.ncbi.nlm.nih.gov/29134598/)

Pope WH, Mavrich TN, Garlena RA, Guerrero-Bustamante CA, Jacobs-Sera D, Montgomery MT, et al., Hatfull. 2017. Bacteriophages of *Gordonia* spp. Display a Spectrum of Diversity and Genetic Relationships. mBio 8: 10.1128/mbio.01069-17. DOI: [10.1128/mbio.01069-17](https://doi.org/10.1128/mbio.01069-17)

Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. Bioinformatics 33(5): 784-786. PubMed ID: [28365761](https://pubmed.ncbi.nlm.nih.gov/28365761/)

Russell DA. 2018. Sequencing, Assembling, and Finishing Complete Bacteriophage Genomes. Methods Mol Biol 1681: 109-125. PubMed ID: [29134591](https://pubmed.ncbi.nlm.nih.gov/29134591/)

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](https://doi.org/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](https://pubmed.ncbi.nlm.nih.gov/38526736/)

Funding: This project has been supported by the Center for Life Sciences Education, The Ohio State University.

Author Contributions: Azzam Almutairi: writing - original draft, investigation. Gabriella Marsala: writing - original draft, investigation. Natalie Mershon: writing - original draft, investigation. Sarah Ball: supervision, investigation, validation, writing - original draft, writing - review editing.

Reviewed By: Anonymous, Randall DeJong

History: Received October 20, 2025 **Revision Received** November 12, 2025 **Accepted** November 17, 2025 **Published Online** November 20, 2025 **Indexed** December 4, 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: Almutairi A, Marsala G, Mershon N, Ball S. 2025. Complete Genome Sequences of seven cluster DE1 phages isolated using *Gordonia terrae*. microPublication Biology. [10.17912/micropub.biology.001909](https://doi.org/10.17912/micropub.biology.001909)