microPublication BIOLOGY

9/15/2026 - *Open Access*

Genome Sequence of Mycobacteriophage Lilbit

Ellie Clavette[1], Andrea Chuva[1,2], Nicholas Gruener[1], Juniper Zinck[1], Sharon Gusky[1§]

[1]STEM, Connecticut State Community College Northwestern, Winsted, CT USA

[2]Torrington High School, Torrington, CT USA

[§]To whom correspondence should be addressed: sharon.gusky@ctstate.edu

Abstract

Mycobacteriophage Lilbit was isolated using *Mycobacterium smegmatis* mc²155. It has a genome consisting of 65,106 base pairs with 110 putative genes and GC content of 63.4%. Based on gene content similarity, it is assigned to actinobacteriophage cluster S. Functions were predicted for 40 genes. There were no tRNAs identified in the genome.

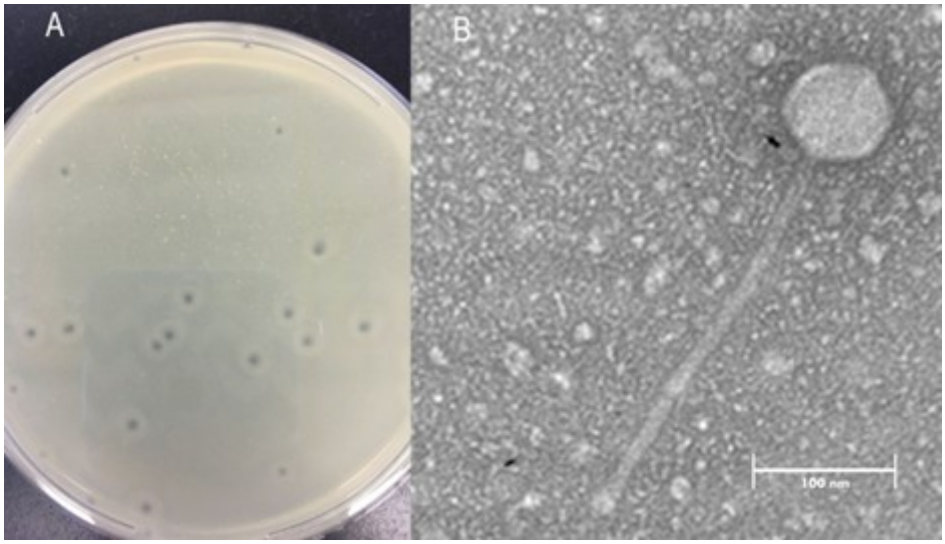

Figure 1. Plaques and Virion Morphology for Lilbit:

Lilbit produces plaques with halos that are ~ 1 mm in diameter (A) and is composed of a capsid 69.81 +/- 3.13 nm in width with a tail 283.58 +/- 7.02 nm in length (B):note the dark specs are artifacts.

Description

Bacteriophages are increasingly being studied for their potential use as therapeutic agents. Mycobacteriophages have been used to treat antibiotic-resistant infections in patients with cystic fibrosis, where *Mycobacterium abscessus* infection can be detrimental to the lungs, as well as in cases of infections by *Mycobacterium avium* and *Mycobacterium chelonae*. (Nick et al. 2022, Dedrick et al. 2023). Here we report on the isolation and characterization of a novel mycobacteriophage, Lilbit, which was isolated using *Mycobacterium smegmatis* mc^{2}155 as a host. There are a number of disease-causing organisms in the Mycobacterium genus including multi-drug resistant strains of *M. tuberculosis* and *M. leprae* which cause tuberculosis and leprosy, respectively. (Garg etal., 2017) Some bacteriophages isolated using *Mycobacterium smegmatis* will be able to infect other members of the genus. (Poxleitner 2018)

Lilbit was isolated from a sample of composting plant material collected in New Haven, Connecticut (GPS coordinates: 41.335200, -72.941760) using an enriched isolation method. The sample was resuspended in 7H9 liquid medium and inoculated with *M. smegmatis* mc^{2}155. After incubation with shaking at 42˚C for several days, the culture was filtered and the filtrate plated in top agar with *M. smegmatis*, resulting in plaques of phage Lilbit, which were purified through multiple rounds of plating until consistent tiny plaques measuring 1mm in size were obtained. Negative stain transmission electron microscopy using uranyl acetate stain revealed Lilbit to possess siphovirus morphology characterized by a non-contractile and flexible tail.

Lilbit's DNA was extracted using the Promega Wizard DNA kit and sequenced by the Pittsburgh Bacteriophage Institute using an Illumina MiSeq (v3 reagents), with libraries prepped using the NEB Ultra II FS kit. This produced 195,422 single-end 150 base raw reads that were assembled using Newbler version 2.9 into a 65,106 base-pair genome with a shot-gun coverage of 419. The assembly and genome termini was checked using Consed V29 (Gordon and Green, 1988), revealing an 11bp 3’ single-strand overhang of 5’-GCGCGCAGCGC at the termini. Lilbit was assigned to cluster S based

on having a gene content similarity of at least 35% to the phages already assigned to cluster S. (Pope et al., 2017). Like other cluster S phages Lilbit has putative HNH endonucleases, putative methyltransferase, and multiple putative glycosyltransferases. (Sevcik et al., 2023)

The sequenced genome was auto-annotated using the Phage Evidence Collection And Annotation Network (PECAAN) v20221109 (Rinehart et al., 2015) using Glimmer v3.02 (Delcher et al., 2007) and GeneMark v4.28 (Lukashin and Borodovsky 1998), with start sites manually refined with Phamerator v606 using Actino draft database v578 (Cresawn etal., 2011) and Starterator v1.2 (Pacey 2016). BLAST, using the Actinobacteriophage BLAST program (Russell and Hatfull 2017) and NCBI non-redundant database v2.2.18 (Altschul et al., 1997) and HHPRED, using the PDB_mmCIF70, Pfam- v36, NCBI Conserved Domains databases (Söding 2005), and NCBIs Conserved Domains databases (Geer et al., 2015), were used to predict the gene functions. DeepTMHMM v1.0 (Chaturvedi 2011) was used to determine if any putative genes coded for transmembrane proteins. Aragorn v1.2.38 (Laslett and Canback 2004) and tRNA scan SE v2.0 (Chan et al., 2021) were used to check for the presence of tRNAs. Default settings were used for all software.

Lilbit was found to encode a total of 110 putative genes, 40 for which putative functions could be assigned. There were 96 genes transcribed in the forward direction, and 14 transcribed in the reverse direction. No tRNAs were found. A programmed -1 translational frameshift was identified for the genes which are predicted to code for tail assembly chaperones. These occur just before a 5,550 bp long gene predicted to encode the tape measure protein. Mycobacteriophage Lilbit is predicted to be lytic based on the absence of identifiable integrase or immunity repressor functions, consistent with other cluster S phages.

Nucleotide sequence accession numbers

Mycobacteriophage Lilbit is available at GenBank with Accession No. PV876982 and Sequence Read Archive (SRA) No. SRX29714289.

Acknowledgements: We would like to acknowledge support from the HHMI Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program and the Pittsburgh Bacteriophage Institute. We acknowledge Shallee Page, Vic Sivanathan, Debbie Jacobs-Sera, Graham Hatfull, and Daniel A. Russell for their support which included genome sequencing and assembly, genome QC, and paper reviews. We also acknowledge Kate Aberger who discovered Lilbit when she was a student at Southern Connecticut State University.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25(17): 3389-402. PubMed ID: 9254694

Chan PP, Lin BY, Mak AJ, Lowe TM. 2021. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Res 49(16): 9077-9096. PubMed ID: 34417604

Chaturvedi N, Shanker S, Singh VK, Sinha D, Pandey PN. 2011.“Hidden markov model for the prediction of transmembrane proteins using MATLAB”, *Bioinformation*, 2011;7(8):418-21. DOI: 10.6026/97320630007418.8

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

Dedrick RM, Guerrero-Bustamante CA, Garlena RA, Russell DA, Ford K, Harris K, et al., Spencer H. 2019. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant Mycobacterium abscessus. Nat Med 25(5): 730-733. PubMed ID: 31068712

Dedrick RM, Smith BE, Cristinziano M, Freeman KG, Jacobs-Sera D, Belessis Y, et al., Hatfull. 2022. Phage Therapy of *Mycobacterium* Infections: Compassionate Use of Phages in 20 Patients With Drug-Resistant Mycobacterial Disease. Clinical Infectious Diseases 76: 103-112. DOI: 10.1093/cid/ciac453

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

Garg K, Gupta R, Bhalla M, Janmeja A. 2017. Multidrug-resistant tuberculosis and leprosy: An unsolved mystery. Lung India 34: 364.doi: 10.4103/lungindia.lungindia_451_15. PMID: 28671168; PMCID: PMC5504894. DOI: 10.4103/lungindia.lungindia_451_15

Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, and Bryant SH, 2015. “CDD: NCBI’s conserved domain database”. *Nucleic Acids Res* 43:D222–D226. DOI: 10.1093/nar/gku1221

Gordon D, Abajian C, Green P. 1998. *Consed:* A Graphical Tool for Sequence Finishing. Genome Research 8: 195-202. https://doi.org/10.1101/gr.8.3.195 PMID: 9521923 DOI: 10.1101/gr.8.3.195

International Committee on Taxonomy of Viruses. 2009 *Siphoviridae*. ICTV. https://ictv.global/report_9th/dsDNA/Siphoviridae

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

Lukashin A. 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Research 26: 1107-1115. DOI: 10.1093/nar/26.4.1107

Nick JA, Dedrick RM, Gray AL, Vladar EK, Smith BE, Freeman KG, et al., Davidson. 2022. Host and pathogen response to bacteriophage engineered against Mycobacterium abscessus lung infection. Cell 185: 1860-1874.e12. DOI: 10.1016/j.cell.2022.04.024

Pacey M. 2016. Starterator guide, University of Pittsburgh, [Online], Available: https://seaphages.org/media/docs/Starterator_Guide_2016.pdf

Pope WH, Mavrich TN, Garlena RA, Guerrero-Bustamante CA, Jacobs-Sera D, Montgomery MT, et al., Hatfull GF. 2017. Bacteriophages of Gordonia spp. Display a Spectrum of Diversity and Genetic Relationships. mBio 8(4): 10.1128/mBio.01069-17. PubMed ID: 28811342

Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull GF. 2018. *HHMI SEA-HAGES phage discovery guide.* Howard Hughes Medical Institute, Chevy Chase, MD. https://seaphagesphagediscoveryguide.helpdocsonline.com/home.

Rinehart C A, Gaffney BL, Smith JR, and Wood J D.2015, PECAAN: Phage Evidence Collection and Annotation Network. Western Kentucky University Bioinformatics and Information Science Center, [Online], Available: https://discover.kbrinsgd.org/login

Russell DA, Hatfull GF. 2016. PhagesDB: the actinobacteriophage database. Bioinformatics 33: 784-786. DOI: 10.1093/bioinformatics/btw711

Sevcik K, Preston P, Aulner M, Noordewier B, Tolsma SS. 2023. Complete Genome Sequences of Cluster S Mycobacteriophages Beelzebub, Raela, and RedRaider77. Microbiology Resource Announcements 12: 10.1128/mra.01173-22. doi: 10.1128/mra.01173-22. Epub 2022 Dec 12. PMID: 36507676; PMCID: PMC9872701. DOI: 10.1128/mra.01173-22

Söding J, Biegert A, and Lupas AN.2005. “The HHpred interactive server for protein homology detection and structure prediction”, *Nucleic Acids Res* 33: W244–W248. DOI: 10.1093/nar/gki408

Funding: This material is based upon work supported by the National Science Foundation under Grants Number 1801062 and Number 2129896. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Conflicts of Interest: The authors declare that there are no conflicts of interest present.

Author Contributions: Ellie Clavette: writing - original draft, writing - review editing, investigation. Andrea Chuva: investigation, writing - original draft. Nicholas Gruener: investigation, writing - original draft. Juniper Zinck: investigation, writing - original draft. Sharon Gusky: data curation, funding acquisition, project administration, resources, supervision, writing - review editing.

Reviewed By: Anonymous, Kristen Butela, Sarah Ball

History: Received September 26, 2025 **Revision Received** April 7, 2026 **Accepted** September 8, 2026 **Published Online** September 15, 2026 **Indexed** September 29, 2026

Citation: Clavette E, Chuva A, Gruener N, Zinck J, Gusky S. 2026. Genome Sequence of Mycobacteriophage Lilbit. microPublication Biology. 10.17912/micropub.biology.001875