

Genomic Characterization of Bacteriophage Babydotz, Isolated from *Microbacterium foliorum*

Harris McFerrin^{1§}, Dahlia Ahmed¹, Justice Austin¹, Classie Baker¹, Skyler Brown¹, Kaionah Cooper¹, Kennedy Drake¹, Liyah Galbert¹, Jillian Harris¹, Nora Herrera¹, Gerald Jones¹, Camerin Kimble¹, Alicia Lewis¹, Leianna Yancey¹, Michael Khalil¹, Shaundessy Perry¹

¹Department of Biology, Xavier University of Louisiana, New Orleans, Louisiana, United States

[§]To whom correspondence should be addressed: hmcferri@xula.edu

Abstract

Microbacteriophage Babydotz was isolated in Moorhead, Minnesota in 2020 by infecting *Microbacterium foliorum* NRRL B-24224. Based on gene content, it is assigned to actinobacteriophage cluster EG. With no identifiable genes involved in lysogeny, it is predicted to have a lytic life cycle.

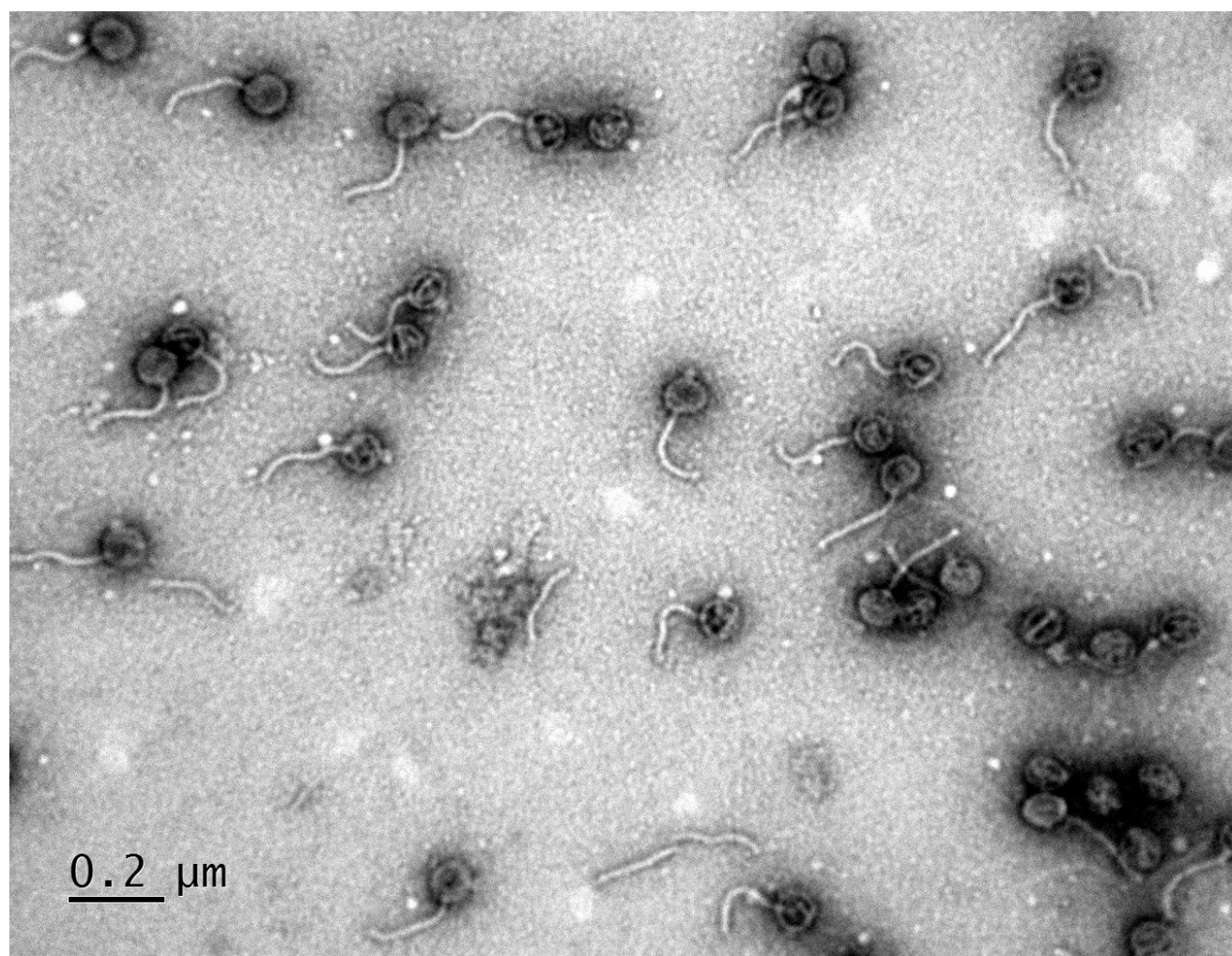


Figure 1. Transmission electron microscopy (TEM) of microbacteriophage Babydotz:

TEM of uranyl acetate-stained microbacteriophage Babydotz shows a siphovirus morphology.

Description

Research on bacteriophages has contributed to our understanding of viruses more broadly and in their development as therapeutics for multi-drug resistant bacterial infections (Kim et al, 2025). Here, we report the genome sequence of Babydotz, a microbacteriophage that infects *Microbacterium foliorum*.

Babydotz was isolated from a soil sample that was collected between two pine trees at the Minnesota State University Moorhead Campus Mall (global positioning system [GPS] 46.8673 N, 96.762 W) when the ambient temperature was

-2°C. Briefly, Babydotz was extracted by washing the soil with PYCa liquid medium, filtering the wash, and plating the filtrate in top agar with *M. foliorum* NRRL B-24224 (Zorowik et al, 2024). After incubation at 30°C for 24 hours, Babydotz formed clear plaques with a diameter of 0.67-0.75mm (n=5), measured using ImageJ software (Schneider, et al, 2012). Negative-staining transmission electron microscopy showed siphovirus morphology with a capsid diameter of 80-90nm (n=5) and tail length of 180-200nm (n=5) (Figure 1).

DNA was extracted from a Babydotz lysate using a Promega Wizard DNA extraction kit prior to sequencing. The Pittsburgh Bacteriophage Institute generated sequencing libraries using the NEB Ultra II FS Kits and sequenced the genome on an Illumina MiSeq using 150-cycle v3 Reagent Cartridges (Russell DA, 2018). Sequencing yielded 613,595 150-base single-end reads. A single bacteriophage contig with 1387-fold coverage was assembled from the raw reads using Newbler v.2.9 and Consed v29 with default parameters (Russell DA, 2018). The resulting genome consisted of 63,076 bp, 217-bp direct terminal repeats and 67.0% G+C nucleotides.

Auto-annotation was performed using DNA Master v5.23.6 (Pope et al., 2018) which employs GeneMark v.3.02 (Delcher et al., 1990) and Glimmer v.2.5p (Borodovsky et al., 2003; Besemer et al., 2005). DNA Master, Phamerator, using Actino_draft database v578 (Cresawn et al., 2011) and PECAAN (Rinehart et al., 2016) were used to manually refine the annotations. Starterator (Pacey, 2014) was used to further evaluate start sites conservation with Glimmer and Genemark. When there was disagreement among these software programs, BLAST results, coding potential and gene length were also used to determine the most likely start for the gene. No tRNA coding genes were identified using tRNAscan-SE v2.0 (Lowe, 2016) or ARAGORN v1.2.41 (Laslett, 2004). Putative functions for 23 of 104 predicted coding ORFs were assigned using BLASTp with an e-value cutoff of 10^{-4} , using the Actinobacteriophage and NCBI non-redundant database (Altschul et al., 1990) and HHPRED, using the PDB_mmCIF70, Pfam- v.36, NCBI Conserved Domains databases with a probability of 90% or greater (Zimmermann et al., 2018; Söding et al., 2005). Nine protein-coding ORFs with membrane-associated domains were determined using SOSUI (Hirokawa et al., 1998) and TMHMM v.1.0.24 (Hallgren et al., 2022). All tools were run with default parameters. Following student genome annotation, the annotation underwent peer review by an experienced SEAPhage-associated faculty and passed quality control inspection. AUG start codon usage was 100%. BabyDotz was assigned to cluster EG based on gene content similarity of at least 35% to phages in the Actinobacteriophage database, phagesdb (<https://phageDB.org>) (Pope et al., 2017; Russell and Hatfull, 2017). Phamerator was used to compare genome architecture and gene synteny between Babydotz and other cluster EG phages. Except for several genes near the left end of the genome, Babydotz exhibits strong synteny with other peer-reviewed cluster EG phages with which there was high sequence alignment (BLASTn E value, Score) including Rowlf (0.0, .93e+04), SallieK (0.0, 1.697e+04), Tissue (0.0, 1.687e+04) and Zagie (0.0, 1.657e+04), consistent with its assignment to cluster EG. Consistent with other cluster EG phages, the central third of the Babydotz genome is transcribed in the forward direction, whereas the flanking regions are transcribed in the reverse direction. An endolysin is encoded immediately upstream of the boundary where transcription switches from the reverse to the forward direction, consistent with the genomic organization observed in other cluster EG phages including Rowlf, SallieK, Tissue and Zagie.

Other predicted genes include those encoding a putative DprA-like DNA processing chain A, three helix-turn-helix DNA binding domains, a terminase, a portal protein, a major capsid protein, a major tail protein, a phosphoesterase, a head-to-tail adaptor, a tape measure protein, three minor tail proteins, an endolysin, an exonuclease, an HNH endonuclease, a RuvC-like resolvase, a MazG-like nucleotide pyrophosphohydrolase, an SSB protein, a DNA primase/helicase and two ribbon-helix-helix DNA binding proteins. No genes encoding immunity repressor or integrase functions could be identified, suggesting that Babydotz is unlikely to establish lysogeny.

Nucleotide sequence accession numbers

Babydotz is available at GenBank with Accession No. OR613478 and Sequence Read Archive (SRA) No. SRR18306109.

Acknowledgements: We thank The Howard Hughes Medical Institute and the Hatfull Lab at the University of Pittsburgh for supporting this research. We would also like to thank Ethan Dotzler and Jonathan Nyandu Kanyinda and their supporting faculty from Minnesota State University Moorhead for the discovery, purification, isolation of DNA from and transmission electron microscopy of phage Babydotz.

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)
- 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)
- Borodovsky M, Mills R, Besemer J, Lomsadze A. 2003. Prokaryotic Gene Prediction Using GeneMark and GeneMark.hmm. *Current Protocols in Bioinformatics* 1: 10.1002/0471250953.bi0405s01. DOI: [10.1002/0471250953.bi0405s01](https://doi.org/10.1002/0471250953.bi0405s01)

- 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 A. 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Research 27: 4636-4641. DOI: [10.1093/nar/27.23.4636](https://doi.org/10.1093/nar/27.23.4636)
- Hallgren J, Tsigos 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](https://doi.org/10.1101/2022.04.08.487609)
- Hirokawa T, Boon-Chieng S, Mitaku S. 1998. SOSUI: classification and secondary structure prediction system for membrane proteins.. Bioinformatics 14: 378-379. DOI: [10.1093/bioinformatics/14.4.378](https://doi.org/10.1093/bioinformatics/14.4.378)
- Kim MK, Suh GA, Cullen GD, Perez Rodriguez S, Dharmaraj T, Chang THW, et al., Sacher JC. 2025. Bacteriophage therapy for multidrug-resistant infections: current technologies and therapeutic approaches. J Clin Invest 135(5): 10.1172/JCI187996. PubMed ID: [40026251](https://pubmed.ncbi.nlm.nih.gov/40026251/)
- 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, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Research 44: W54-W57. DOI: [10.1093/nar/gkw413](https://doi.org/10.1093/nar/gkw413)
- Pacey, M. Starterator Guide. (2014) https://phagesdb.org/media/docs/Starterator_Guide_2014_2.pdf
- Pope WH, Jacobs-Sera D. 2017. Annotation of Bacteriophage Genome Sequences Using DNA Master: An Overview. Methods in Molecular Biology, Bacteriophages : 217-229. DOI: [10.1007/978-1-4939-7343-9_16](https://doi.org/10.1007/978-1-4939-7343-9_16)
- 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](https://pubmed.ncbi.nlm.nih.gov/28811342/)
- Rinehart CA, Gaffney B, Wood JD, Smith J. PECAAN, a Phage Evidence Collection And Annotation Network. (2016) <https://discover.kbrinsgd.org/login>
- Russell DA, Hatfull GF. 2016. PhagesDB: the actinobacteriophage database. Bioinformatics 33: 784-786. DOI: [10.1093/bioinformatics/btw711](https://doi.org/10.1093/bioinformatics/btw711)
- 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](https://doi.org/10.1007/978-1-4939-7343-9_9)
- Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. Nature Methods 9: 671-675. DOI: [doi:10.1038/nmeth.2089](https://doi.org/10.1038/nmeth.2089)
- 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)
- Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, et al., Alva. 2018. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. Journal of Molecular Biology 430: 2237-2243. DOI: [10.1016/j.jmb.2017.12.007](https://doi.org/10.1016/j.jmb.2017.12.007)
- Zorawik M, Jacobs-Sera D, Freise AC, SEA-PHAGES, Reddi K. 2024. Isolation of Bacteriophages on Actinobacteria Hosts. Methods in Molecular Biology, Phage Engineering and Analysis : 273-298. DOI: [10.1007/978-1-0716-3798-2_17](https://doi.org/10.1007/978-1-0716-3798-2_17)

Funding: Funds supporting this research were provided by the J.W. Carmichael Endowed Professorship from the Louisiana Board of Regents and Xavier University of Louisiana.

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

Author Contributions: Harris McFerrin: conceptualization, project administration, writing - original draft, writing - review editing, data curation, formal analysis. Dahlia Ahmed: writing - original draft, investigation, data curation. Justice Austin: investigation. Classie Baker: investigation. Skyler Brown: investigation. Kaionah Cooper: investigation. Kennedy Drake: investigation. Liyah Galbert: investigation. Jillian Harris: investigation. Nora Herrera: investigation. Gerald Jones: investigation. Camerin Kimble: investigation. Alicia Lewis: investigation. Leianna Yancey: investigation. Michael Khalil: investigation. Shaundessy Perry: investigation.

Reviewed By: Kurt Regner, Sarah Ball

History: Received September 23, 2025 **Revision Received** July 14, 2026 **Accepted** August 7, 2026 **Published Online** August 10, 2026 **Indexed** August 24, 2026

Copyright: © 2026 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: McFerrin H, Ahmed D, Austin J, Baker C, Brown S, Cooper K, et al., Perry S. 2026. Genomic Characterization of Bacteriophage Babydotz, Isolated from *Microbacterium foliorum*. microPublication Biology. [10.17912/micropub.biology.001869](https://doi.org/10.17912/micropub.biology.001869)