

# Genome Sequence of the *Mycobacterium smegmatis* Bacteriophage Eugenia

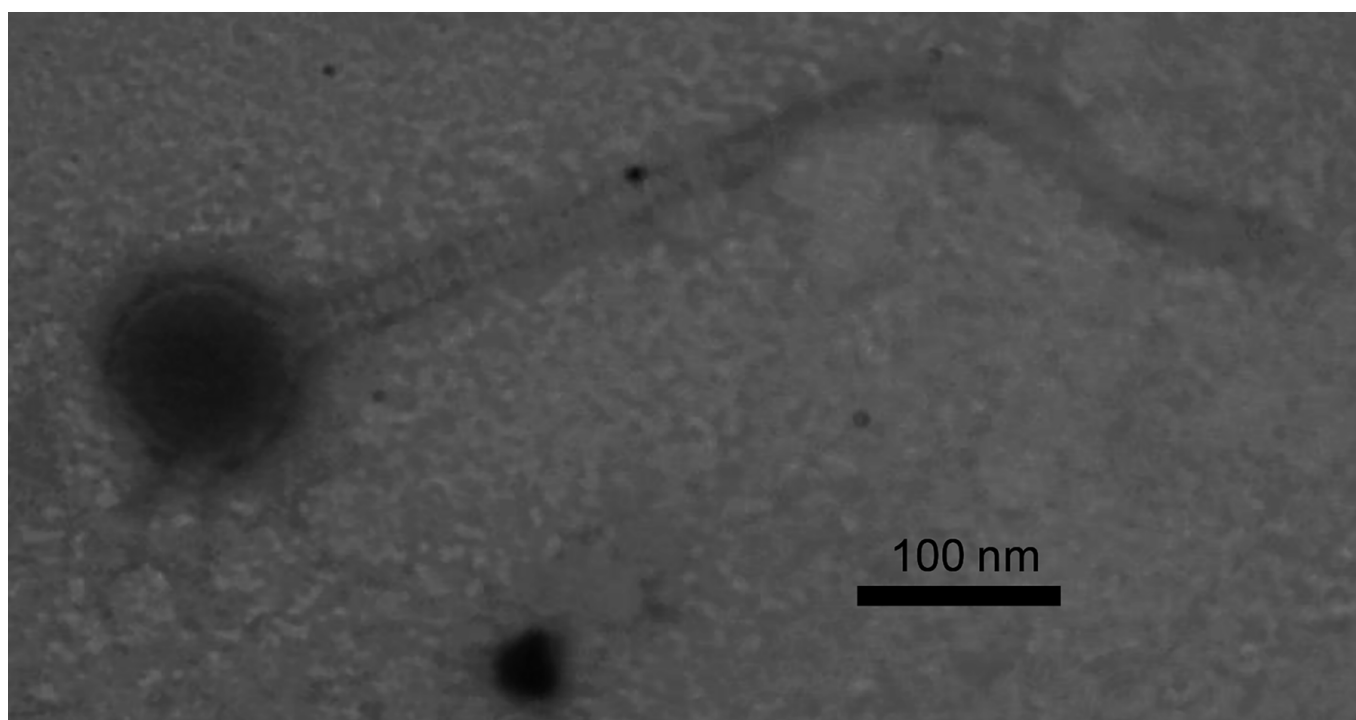
Vipaporn Phuntumart<sup>1§</sup>, Lucia Boulos<sup>1</sup>, Bella Nunnally<sup>1</sup>, Isabella Lima<sup>1</sup>, John Motter<sup>1</sup>, Olivia Sidoti<sup>1</sup>, Sam Rutherford<sup>1</sup>, Hsin-Ho Wei<sup>1</sup>, Raymond Larsen<sup>1</sup>, Jill H Zeilstra-Ryalls<sup>1</sup>

<sup>1</sup>Biological Sciences, Bowling Green State University, Bowling Green, Ohio, United States

<sup>§</sup>To whom correspondence should be addressed: vphuntu@bgsu.edu

## Abstract

We report the discovery and genome sequence of mycobacteriophage Eugenia, isolated from soil samples collected in Akron, OH. Eugenia is a double-stranded DNA virus with a genome size of 69,139 bp, featuring 104 predicted protein-encoding genes, with 32 of these genes assigned putative functions.



**Figure 1. Transmission electron micrograph of Eugenia taken using Zeiss EM10:**

The phage was negatively stained with 2% uranyl acetate. Its capsid diameter is 75 nm and the tail length is 530 nm.

## Description

Bacteriophages have fueled fundamental discoveries in biology, the development of biotechnological tools, and therapeutics. The latter is highlighted by the recent treatment of a patient infected with multidrug-resistant *Mycobacterium abscessus* using a cocktail of phages that were isolated from the nonpathogenic bacterium *Mycobacterium smegmatis* mc<sup>2</sup> 155 (Dedrick et al., 2019; Mohan et al., 2015). Here, we present the characterization and sequence of a new mycobacteriophage, Eugenia. This phage was isolated from soil in a flower bed in Akron, Ohio (41.1871 N, 81.6822 W), using an enrichment procedure (Zorawik et al., 2024). Briefly, the soil sample was washed in nutrient broth (7H9), the wash was filtered using a 0.22  $\mu$ m pore size filter, and the filtrate was inoculated with *M. smegmatis* mc<sup>2</sup> 155. Following incubation in 7H9 media for 48 hr at 37°C, the bacteria were removed by filtration, and the filtrate was plated in top agar with *M. smegmatis*. After two days of incubation at 37°C, Eugenia phage were isolated from a clear plaque with a diameter of approximately 1 mm. Eugenia was then plaque-purified through five rounds of plating. Negative stain (2% uranyl acetate) transmission electron microscopy (Zeiss EM10) revealed Eugenia to have a siphovirus morphology with a capsid size of 75.2  $\pm$  9 nm and tail length of 445  $\pm$  57 nm.

Genomic DNA was extracted from a high titer lysate using the Promega Wizard DNA cleanup kit, and a library was generated using the Illumina Truseq Nano DNA prep kit. Sequencing was performed using a MiSeq System and v3 600 Cycle Reagent kit (2 x 300 paired end reads), yielding ~ 2,280,377 bp of 300-base single-end reads. The raw reads were assembled using Newbler v2.9 and default parameters achieving a coverage of 4,682 (Margulies et al., 2005). Methods described by Russell (2018) were used to ensure the completeness and accuracy of the sequence.

The *Eugenia* genome was annotated using bioinformatic tools and databases that included DNAMaster v5.23.6 (Pope and Jacobs-Sera, 2018) embedded with Glimmer v3.02b (Delcher et al., 1999) and Genemark v2.5 (Besemer and Borodovsky, 2005), PECAAN v1.0 (Rinehart et al., 2016; <https://discover.kbrinsgd.org/>), PhagesDB BLAST: Actinobacteriophage database (Russell and Hatfull, 2017), NCBI BLAST: non-redundant database (Altschul et al., 1990), Phamerator database v505 (Cresawn et al., 2011), Starterator database v505 (Pacey, 2016; <http://phages.wustl.edu/starterator/>), HHPRED v3.2.0: PDB mmCIF70, Pfam-A (Söding et al., 2005), NCBI Conserved Domain databases (Wang et al., 2022), Aragorn v1.1 and v1.2.38 (Laslett and Canback, 2004), tRNAscanSE v2.0 16. (Lowe and Eddy, 1997), TMHMM v2.0 (Krogh et al., 2001), and SOSUI v1.11 (Hirokawa et al., 1998). *Eugenia* was assigned to cluster B, subcluster B1 based on gene content similarity of at least 35% to phages in the Actinobacteriophage Database (Russell and Hatfull, 2017). Its genome is 69,139 bp, making it among the largest genomes of B1 subcluster phages, to date. It has 3' single-stranded overhang of 9 bases (CGCGGGGGA) and a G+C content of 66.5%, which is similar to that of the host *M. smegmatis* mc<sup>2</sup> 155 (67.4%) bacterium (Mohan et al., 2015). A total of 104 putative protein-encoding genes were identified, and of these 32 were assigned putative functions, without any tRNA. The DNA sequence of *Eugenia* is available at [GenBank PP978882.1](https://www.ncbi.nlm.nih.gov/genbank/PP978882.1) and the raw sequence data have been deposited in the Sequence Read Archive (SRA) under No. [SRX25258414](https://www.ncbi.nlm.nih.gov/sra/SRX25258414).

Phages of the B1 subcluster have therapeutic potential because they display inhibitory action toward biofilm formation and form plaques on strain 4XR1, an isoniazid-resistant derivative of *M. smegmatis* mc<sup>2</sup> 155 (Mohan et al., 2015). While *Eugenia* forms clear plaques and is predicted to be lytic based on the lack of identifiable immunity repressor or integrase functions, phage-release experiments suggest that some cluster B1 phages may be temperate (Das et al., 2024). A RepA-like protein in B1 phages that is also present in *Eugenia* (gp58) has been predicted to be involved in lysogeny (Das et al., 2024). However, this protein bears no similarity to RepA-like proteins of mycobacteriophages whose function has been experimentally linked to lysogeny (Wetzel et al., 2020). Nevertheless, it raises important questions about the mechanisms of action and potential applications of *Eugenia* in therapeutic contexts, warranting further investigation into its biology.

**Acknowledgements:** We would like to thank all the members of the SEA-PHAGES SMART team for verifying the annotation of *Eugenia* genome, and the SEA Team for their critical review of this manuscript. We also thank Dr. Marilyn Cayer for her assistance with the transmission electron microscope used to capture images of the phage.

## References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215(3): 403-10. PubMed ID: [2231712](https://pubmed.ncbi.nlm.nih.gov/2231712/)
- 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/)
- Das R, Arora R, Nadar K, Saroj S, Singh AK, Patil SA, et al., Bajpai U. 2024. Insights into the genomic features and lifestyle of B1 subcluster mycobacteriophages. *J Basic Microbiol* 64(6): e2400027. PubMed ID: [38548701](https://pubmed.ncbi.nlm.nih.gov/38548701/)
- 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](https://pubmed.ncbi.nlm.nih.gov/31068712/)
- Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res* 27(23): 4636-41. PubMed ID: [10556321](https://pubmed.ncbi.nlm.nih.gov/10556321/)
- Hirokawa T, Boon-Chieng S, Mitaku S. 1998. SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* 14(4): 378-9. PubMed ID: [9632836](https://pubmed.ncbi.nlm.nih.gov/9632836/)
- Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305(3): 567-80. PubMed ID: [11152613](https://pubmed.ncbi.nlm.nih.gov/11152613/)

12/14/2024 - Open Access

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](#)

Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25(5): 955-64. PubMed ID: [9023104](#)

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bembien LA, et al., Rothberg JM. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437(7057): 376-80. PubMed ID: [16056220](#)

Mohan A, Padiadpu J, Baloni P, Chandra N. 2015. Complete Genome Sequences of a *Mycobacterium smegmatis* Laboratory Strain (MC2 155) and Isoniazid-Resistant (4XR1/R2) Mutant Strains. *Genome Announc* 3(1). PubMed ID: [25657281](#)

Pacey, M. 2016. Starterator guide. Pittsburgh, PA University of Pittsburgh.

Pope WH, Jacobs-Sera D. Annotation of bacteriophage genome sequences using DNA master: an overview. In: Clokie MRJ, Kropinski AM, Lavigne R, editors. *Bacteriophages: methods and protocols*. Vol. 3. New York: Springer; 2018. p. 217–229.

Rinehart, CA, Gaffney, BL, Smith, JR, & Wood, JD. 2016. PECAAN: Phage Evidence Collection and Annotation Network user guide. Western Kentucky University Bioinformatics and Information Science Center, Bowling Green, KY.

Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. *Bioinformatics* 33(5): 784-786. PubMed ID: [28365761](#)

Russell DA. Sequencing, assembling, and finishing complete bacteriophage genomes. In: Clokie MRJ, Kropinski AM, Lavigne R, editors. *Bacteriophages: methods and protocols*. Vol. 3. New York: Springer; 2018. p. 109–125.

Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33(Web Server issue): W244-8. PubMed ID: [15980461](#)

Wang J, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu S, et al., Marchler-Bauer A. 2023. The conserved domain database in 2023. *Nucleic Acids Res* 51(D1): D384-D388. PubMed ID: [36477806](#)

Wetzel KS, Aull HG, Zack KM, Garlena RA, Hatfull GF. 2020. Protein-Mediated and RNA-Based Origins of Replication of Extrachromosomal Mycobacterial Prophages. *mBio* 11(2). PubMed ID: [32209683](#)

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](#)

**Funding:** Support from this research was provided by the Department Biological Sciences at Bowling Green State University, the SEA-PHAGES Program, and the University of Pittsburgh.

**Author Contributions:** Vipaporn Phuntumart: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, writing - original draft, writing - review editing. Lucia Boulos: methodology. Bella Nunnally: formal analysis. Isabella Lima: formal analysis. John Motter: formal analysis. Olivia Sidoti: formal analysis. Sam Rutherford: formal analysis. Hsin-Ho Wei: data curation, validation. Raymond Larsen: investigation, methodology, supervision. Jill H Zeilstra-Ryalls: investigation, methodology, supervision, writing - review editing, funding acquisition.

**Reviewed By:** Kristen Butela, Anonymous

**History:** Received October 28, 2024 **Revision Received** December 9, 2024 **Accepted** December 13, 2024 **Published Online** December 14, 2024 **Indexed** December 28, 2024

**Copyright:** © 2024 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:** Phuntumart, V; Boulos, L; Nunnally, B; Lima, I; Motter, J; Sidoti, O; et al.; Zeilstra-Ryalls, JH (2024). Genome Sequence of the *Mycobacterium smegmatis* Bacteriophage Eugenia. *microPublication Biology*. [10.17912/micropub.biology.001401](#)