

Genome Sequence of the *Gordonia rubripertincta* Phage Yucky (Cluster CT)

Caitlin Carter^{1*}, Paxton R. Evitts^{1*}, Cheyenne Helton^{1*}, Madeline F. Kaiser^{2*}, Jehee F. Lee^{1*}, Mackenzie M. Peters^{1*}, Cade Riley^{1*}, Mackenzie L. Sanders^{1*}, Chad Walker^{1*}, Danielle K. Watt¹, Pamela L. Connerly¹, Elizabeth E. Rueschhoff^{1§}

¹School of Natural Sciences, Indiana University Southeast, New Albany, IN, US

²School of Social Sciences, Indiana University Southeast, New Albany, IN, US

§To whom correspondence should be addressed: erueschh@iu.edu

*These authors contributed equally.

Abstract

Novel bacteriophage Yucky infects *Gordonia rubripertincta*, a bacterium known for its bioremediation potential. Phage particles have a siphovirus morphology and form 1.2 mm ± 0.5 mm-wide plaques (n=10). The genome of Yucky is 47,803 bp in length and encodes 74 protein-coding genes. The phage was assigned to cluster CT based on gene content similarity of at least 35% to actinobacteriophages. Thirty-nine gene products have putative functions, including genes involved in lysis, such as two lysin A domain genes as well as a lysin B gene.

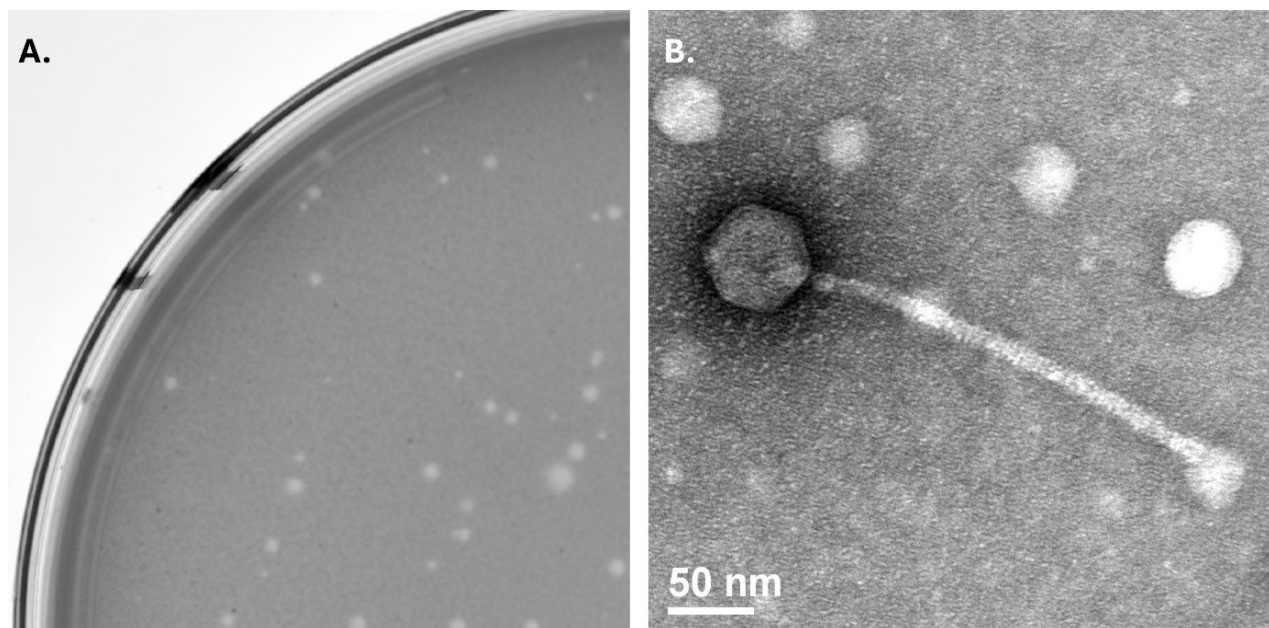


Figure 1. Plaque and virion morphology.:

A. Plaque Morphology of *Gordonia* Phage Yucky. The plaque size of Yucky is variable, but average approximately 1.2 mm ± 0.5 mm in size (n=10). Shown in a segment of a standard 100 mm plate. B. Transmission Electron Micrograph of Yucky. Negative staining with uranyl acetate (2%) reveals phage particles with a capsid diameter of 63.7 nm ± 2.6 nm and a tail length of 270.6 nm ± 17.8 nm (n=5). Measurements of plaque diameter, capsids, and tails were made utilizing ImageJ software (v1.54k) (Schneider et al. 2012).

Description

Strains of *Gordonia rubripertincta* have been characterized as bioremediators of various environmental pollutants. For example, *G. rubripertincta* strain 112 degrades aromatic and aliphatic compounds (Frantsuzova et al. 2023), while *G. rubripertincta* CWB2 degrades styrene (Lienkamp et al. 2021). Identifying and studying bacteriophages that infect these bioremediators may be informative in controlling and utilizing these bacteria in the environment.

Bacteriophage Yucky was isolated from soil collected in Jeffersonville, IN, USA (Global Positioning Coordinates [GPS]: 38.35299 N, 85.71948 W). A mixture of soil and PYCa media was incubated for approximately 2 hours at 26 °C and shaken at approximately 250 RPM. The mixture was then centrifuged, and the supernatant was filtered using a 0.2-μm filter. The filtrate was inoculated with *G. rubripertincta* NRRL B-16540 to enrich for phages capable of infecting this host and incubated for approximately 5 days at 220 rpm at 26 degrees Celsius. The sample was again filtered (0.2 μm pore

size) and plated using soft agar overlay. Plates were incubated for approximately 48 hours at 26 °C. Plaques formed were round and clear, measuring 1.2 mm ± 0.5 mm in diameter (n=10). Yucky was further purified through two rounds of plating before high titer lysates were collected (Zorawik et al. 2024). Negative stain electron microscopy (2% uranyl acetate) revealed phage particles with a siphovirus morphology, with a capsid size of 63.7 nm ± 2.6 nm and a tail length of 270.6 nm ± 17.8 nm (n=5).

DNA was isolated from a high titer lysate of Yucky using a QIAGEN DNeasy Blood and Tissue Kit (Jakociune and Moodley 2018). The DNA was prepared for sequencing using a New England Biolabs Ultra II FS Library Kit and sequenced with an Illumina NextSeq 1000 sequencer (X-LEAP PI Kit). Raw 100 base reads were trimmed with cutadapt 4.7 (using the option: `-nextseq-trim 30`) and filtered with skewer 0.2.2 (using the options: `-q 20 -Q 30 -n -l 50`) prior to assembly (Martin 2011; Jiang et al. 2014; Wick et al. 2017; Gordon et al. 1998). Newbler v2.9 (Miller et al. 2010) and Consed v29 (Gordon and Green 2013) were used to assemble the trimmed reads and check them for completeness (Russell 2018), with an assembly coverage of approximately 3577x. The genome is 47,803-bp in length, has a GC content of 60.5%, and has genome ends characterized by 3' single-stranded overhangs 10 bases in length (5'-CGGTAGGCTT-3').

The genome of Yucky was auto-annotated utilizing Glimmer v3.02b and Genemark v2.5p (Delcher et al. 2007; Besemer and Borodovsky 2005). Manual annotation was performed using the DNA Master program v5.23.6 build 2705 (Pope and Jacobs Sera 2018; <http://cobamide2.bio.pitt.edu>), BLASTp v2.16.0 utilizing the NCBI nonredundant and Actinobacteriophage databases (Altschul et al. 1990), Starterator v594 (<http://phages.wustl.edu/starterator/>), Phamerator utilizing Actino_Draft v594 (Cresawn et al. 2011), HHPred utilizing the PDB_mmCIF70_30_Mar, Pfam-A v37, NCBI Conserved Domains databases (CD) v3.19, and UniProt-SwissProt-viral70_3_Nov2021 databases (Söding et al. 2005), Aragorn v1.2.41 (Laslett and Canback 2004), and deepTMHMM v1.0.42 (Krogh et al. 2001). Default settings were used for all software.

Seventy-four total protein-coding genes were predicted, and no tRNA genes were identified. Of these 74 genes, 39 were assigned putative functions and 35 were hypothetical proteins. Based on gene content similarity of at least 35% to phages in the Actinobacteriophage database, Yucky was assigned to cluster CT (Russell and Hatfull 2017; Pope et al 2017). As a member of cluster CT, Yucky contains genes assigned to a majority of common functions found in all cluster CT phages (Mitchell et al. 2024) including viral structure and assembly, DNA replication and recombination, and cell lysis. The domains of the endolysin are split into two genes: one encoding the L-Ala-D-Glu peptidase domain (gene 20) and the other encoding the glycosyl hydrolase domain (gene 21). In addition, a lysin B gene was identified (gene 49). Yucky also contains genes less commonly found within cluster CT phages including a predicted PAPS reductase-like domain (gene 2, found in only 4 cluster CT phages), and a predicted SprT-like protease (gene 37, found in 18 cluster CT phages).

To date, Yucky is one of 77 sequenced phages in cluster CT. Within the cluster, Yucky is most similar to phages Vine and PotPie, with a gene content similarity (GCS) of 93.2% and 93.8%, respectively (Russell and Hatfull 2017). Relative to one another, each phage encodes short segments of unique genes; Vine genes 1, 3, 4, 35, and 46, PotPie genes 29, 30, and 33, and Yucky genes 34, 35, and 46 are all unique. No putative genes with immunity repressor, integrase or DNA partitioning functions were identified, suggesting that Yucky is unlikely to develop lysogeny.

Nucleotide sequence accession numbers for Yucky is available at GenBank with Accession No. [PV876943](https://www.ncbi.nlm.nih.gov/nuccore/PV876943) and Sequence Read Archive (SRA) No. [SRX28484036](https://www.ncbi.nlm.nih.gov/sra/SRX28484036).

Acknowledgements: We gratefully acknowledge Shelia Gustafson, at Indiana University Southeast, who provided the soil sample from which Yucky was isolated. We also acknowledge the combined efforts of the 2024-2025 SEA-PHAGES cohort at Indiana University Southeast. Additional thanks to Dr. Barry Stein, who imaged the bacteriophage at the Indiana University Electron Microscopy Center and Daniel A. Russell at the University of Pittsburgh, who sequenced and aligned the genome. Not to be forgotten are the SEA-PHAGES program and the HHMI for their continued support and assistance.

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403-410. 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/)
- Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23: 673-679. PubMed ID: [17237039](https://pubmed.ncbi.nlm.nih.gov/17237039/)
- Frantsuzova E, Bogun A, Solomentsev V, Vetrova A, Streletskii R, Solyanikova I, Delegan Y. 2023. Whole Genome Analysis and Assessment of the Metabolic Potential of *Gordonia rubripertincta* Strain 112, a Degrader of Aromatic and Aliphatic Compounds. *Biology* 12: 721. PubMed ID: [37237534](https://pubmed.ncbi.nlm.nih.gov/37237534/)

- Gordon D, Abajian C, Green P. 1998. *Consed*: A Graphical Tool for Sequence Finishing. *Genome Research* 8: 195-202. PubMed ID: [9521923](#)
- Gordon D, Green P. 2013. *Consed*: a graphical editor for next-generation sequencing. *Bioinformatics* 29: 2936-2937. PubMed ID: [23995391](#)
- Jakočiūnė Di, Moodley A. 2018. A Rapid Bacteriophage DNA Extraction Method. *Methods and Protocols* 1: 27. PubMed ID: [31164569](#)
- Jiang H, Lei R, Ding SW, Zhu S. 2014. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* 15: 10.1186/1471-2105-15-182. PubMed ID: [24925680](#)
- 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](#)
- 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](#)
- Lienkamp AC, Burnik J, Heine T, Hofmann E, Tischler D. 2021. Characterization of the Glutathione S-Transferases Involved in Styrene Degradation in *Gordonia rubripertincta* CWB2. *Microbiol Spectr* 9(1): e0047421. PubMed ID: [34319142](#)
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 17:10. DOI: [10.14806/ej.17.1.200](#)
- Miller JR, Koren S, Sutton G. 2010. Assembly algorithms for next-generation sequencing data. *Genomics* 95: 315-327. PubMed ID: [20211242](#)
- Mitchell JC, Fogarty MP, Pollenz RS. 2024. Cluster CT annotation report. HHMI Science Education Alliance (SEA) Faculty Group, QUBES Educational Resources. DOI: [10.25334/ZVY0-X671](#)
- 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](#)
- Pope WH, Montgomery MT, Bonilla JA, Dejong R, Garlena RA, Guerrero Bustamante C, et al., Hatfull. 2017. Complete Genome Sequences of 38 *Gordonia* sp. Bacteriophages. *Genome Announcements* 5: 10.1128/genomea.01143-16. PubMed ID: [28057748](#)
- Russell DA. 2018. Sequencing, Assembling, and Finishing Complete Bacteriophage Genomes. *Methods Mol Biol* 1681: 109-125. PubMed ID: [29134591](#)
- Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. *Bioinformatics* 33(5): 784-786. PubMed ID: [28365761](#)
- Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9: 671-675. PubMed ID: [22930834](#)
- Soding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research* 33: W244-W248. PubMed ID: [15980461](#)
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Computational Biology* 13: e1005595. PubMed ID: [28594827](#)
- 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: Funding for this project was provided by Indiana University Southeast and the James Y. McCullough Memorial Endowment.

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

Author Contributions: Caitlin Carter: investigation. Paxton R. Evitts: investigation. Cheyenne Helton: investigation. Madeline F. Kaiser: investigation. Jehée F. Lee: investigation. Mackenzie M. Peters: investigation. Cade Riley: investigation. Mackenzie L. Sanders: investigation. Chad Walker: investigation. Danielle K. Watt: data curation, investigation, validation. Pamela L. Connerly: data curation, investigation, project administration, supervision, validation, writing - original draft. Elizabeth E. Rueschhoff: data curation, investigation, supervision, validation, writing - original draft, project administration.

Reviewed By: Anonymous

History: Received December 15, 2025 **Revision Received** April 2, 2026 **Accepted** April 14, 2026 **Published Online** April 16, 2026 **Indexed** April 30, 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: Carter C, Evitts PR, Helton C, Kaiser MF, Lee JF, Peters MM, et al., Rueschhoff EE. 2026. Genome Sequence of the *Gordonia rubripertincta* Phage Yucky (Cluster CT). microPublication Biology. [10.17912/micropub.biology.001981](https://doi.org/10.17912/micropub.biology.001981)