

Complete annotated genome sequence of *Microbacterium paraoxydans* phage Evcara, a cluster GI podovirus isolated from compost.

Hannah B. Bonogafsky¹, Kelly M. Freer¹, Taylor B. Sanderson¹, Kira A. Yost¹, Ayden J. Pyle¹, Ginavieve Rowley¹, Maria D. Gainey^{1§}

¹Department of Chemistry and Physics, Western Carolina University, Cullowhee, North Carolina, US

[§]To whom correspondence should be addressed: mdgainey@email.wcu.edu

Abstract

Bacteriophage Evcara is a podovirus isolated on *Microbacterium paraoxydans* NRRL B-24275. Its genome is 16,285 bp in length and contains 22 predicted protein-coding genes. Evcara, has been assigned to cluster GI with *Microbacterium foliorum* phages PineapplePizza and Curie that share 10 homologues with the well-characterized *Bacillus subtilis* phage phi29.

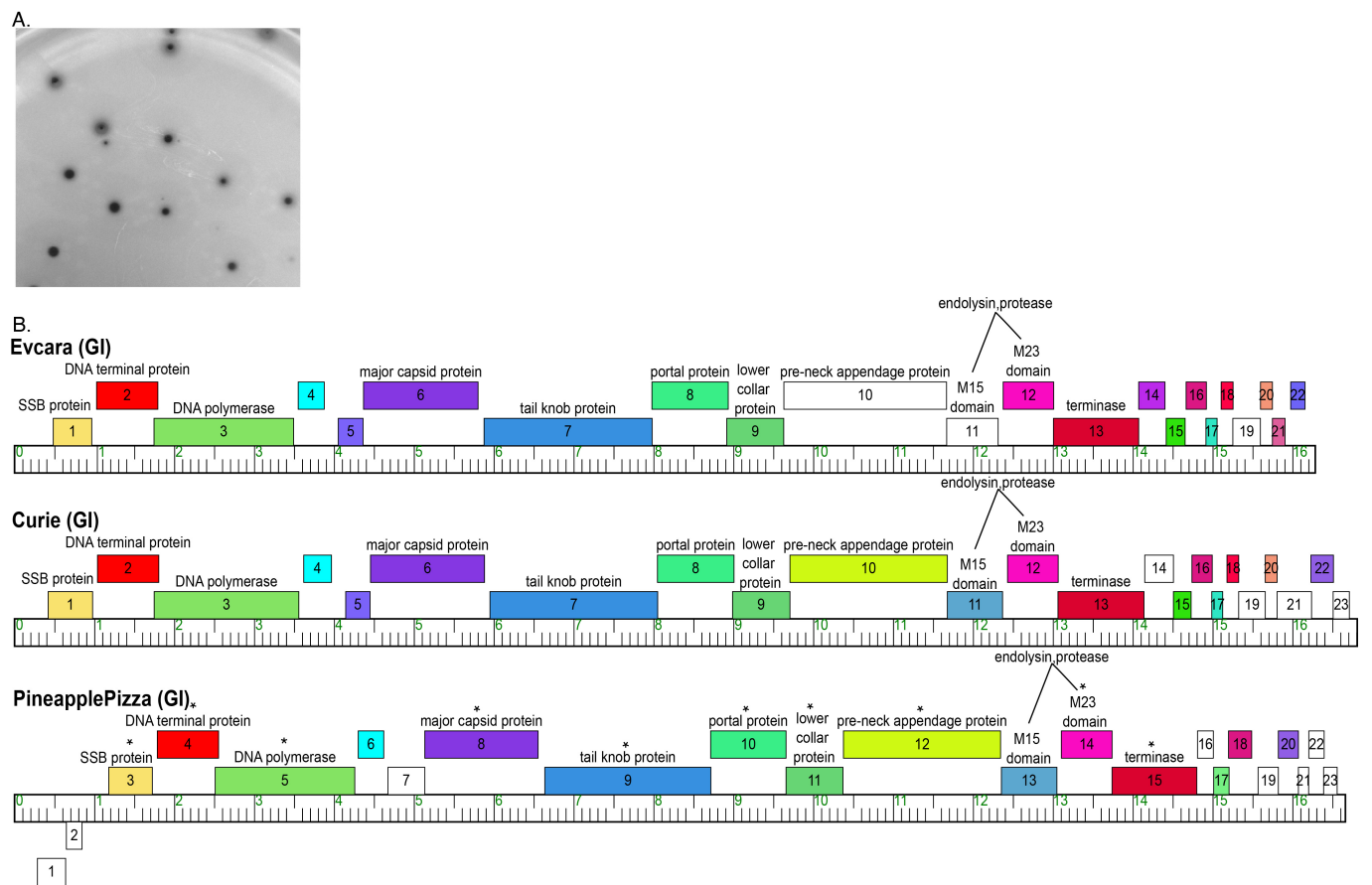


Figure 1. Evcara plaque morphology and cluster GI genomic organization:

(A) Evcara forms plaques with clear centers and two halos that vary in size (1mm to 6mm in diameter) and halo appearance when plated on *M. paraoxydans* PYCa top agar overlays incubated for 48hr at 30°C. (B) A representation of the genomes of all three cluster GI bacteriophages was generated using Phamerator. Genes are represented with numbered boxes, and assigned functions are listed. Boxes above and below the ruler represent genes that are transcribed rightwards and leftwards, respectively. PineapplePizza genes with phi29 homologues are denoted with a “*.” Genes represented as white boxes are genes with no homologues in the Actinobacteriophage database.

Description

To date, the Actinobacteriophage database (phagesDB.org), hosts genome sequences for 690 *Microbacterium* phages, isolated using 16 species of *Microbacterium* (Russell & Hatfull, 2017). These phages exhibit considerable genetic diversity, are predominantly lytic, and have been classified into 23 genetic clusters (Jacobs-Sera et al., 2020; Pope et al., 2017). Here we describe Evcara, a phage isolated using *Microbacterium paraoxydans* NRRL B-24275, one of only three cluster GI bacteriophages isolated, to date.

In the fall of 2023, a sample from the Cullowhee Community Garden compost pile in Cullowhee, North Carolina, USA (GPS 35.317778 N, 83.180833 W) was washed with peptone-yeast calcium (PYCa) media. The wash was filtered (0.22µm) and then enriched by inoculating the filtrate with *M. paraoxydans*. After incubation with shaking (48hrs at 30°C), the culture was filtered and the filtrate spotted onto PYCa plates containing top agar overlays supplemented with *M. paraoxydans*. After incubation at 30°C for 48h, a clearance with a turbid center was observed, indicative of bacteriophage infection. Bacteriophage Evcara was then plated to obtain individual plaques and purified via 5 additional rounds of plating, after which a lysate was prepared (Zorawik et al., 2024). Evcara plaques vary in size (1mm to 6mm in diameter) and have clear centers encircled by two halos of varying turbidity (Fig. 1A).

Genomic DNA was isolated from the lysate using a Promega Wizard DNA cleanup kit and prepared for sequencing using an NEB Ultra II library kit. A total of 561,542, 150bp single-end reads (5,008x genome coverage) were generated using an Illumina MiSeq instrument (v3 reagents) at the Pittsburgh Bacteriophage Institute. Raw reads were assembled into a single contig using Newbler v2.9, and checked for accuracy using Consed v.29 (Gordon & Green, 2013; Margulies et al., 2005; Russell, 2018). Default parameters were used for all software unless otherwise specified.

The Evcara genome is 16,285 bp in length (54.0% GC) with 121bp inverted repeats at the genome ends that are likely covalently bound by terminal proteins as described for phi29 (Yehle, 1978). Utilizing DNA Master v2705 (<http://cobamide2.bio.pitt.edu/>), embedded with Genemark v2.5 p and Glimmer v3.02, an automated annotation of the Evcara Genome was generated (Besemer & Borodovsky, 2005; Delcher et al., 2007). Start sites were manually reviewed using Starterator v1.2 (<http://phages.wustl.edu/starterator/>), Phamerator v567 (Cresawn et al., 2011) using the Actino_draft v578, and PECAAN v.20240320 (<https://discover.kbrinsgd.org/evidence/summary>). BLASTp (Altschul et al., 1990) searches against the NCBI non-redundant and Actinobacteriophage databases, and HHPred (Söding et al., 2005) searches against the PDB_mmCIF70_8_Mar, Pfam-A_v37, UniProt-SwissProt-viral70_3, and NCBI_Conserved_Domains(CD)_v3.19 databases were used to assign putative functions. No tRNA genes were identified using Aragorn v1.2.41 (Laslett & Canback, 2004) or tRNAscanSE v.2.0 (Lowe & Chan, 2016).

Whole-genome alignment using NCBI BLASTn (Altschul et al., 1990) showed no significant nucleotide similarity with any other sequences in NCBI's core nucleotide blast database, with the exception of a 233bp region (81% identity) to *Microbacterium* phage Curie's genome. However, genome annotation revealed extensive shared gene content with *Microbacterium foliorum* podoviruses Curie and PineapplePizza, surpassing the clustering threshold of 35% gene content similarity allowing Evcara to be assigned to actinobacteriophage cluster GI (Pope et al., 2017). Despite their lack of nucleotide similarity, the 3 cluster GI bacteriophages have a highly conserved genome architecture consisting of 22 to 23 predicted protein coding genes all transcribed in the same direction, with the exception of PineapplePizza's first two genes which are transcribed in the opposite direction (Fig. 1B). Evcara has 22 genes, 11 of which were assigned a function. These functions are conserved for all members of cluster GI, 10 of which were previously reported to be homologous to the well-studied podovirus *Bacillus* phage phi29 (Horton et al., 2024; LaBianca et al., 2023; Xu et al., 2019).

Nucleotide sequence accession numbers

GenBank Accession No.: [PQ114753](https://www.ncbi.nlm.nih.gov/nuccore/PQ114753)

Sequence Read Archive (SRA) No.: [SRX25734226](https://www.ncbi.nlm.nih.gov/sra/SRX25734226)

Acknowledgements: We would like to thank Dan Russel for sequencing and assembly of the Evcara genome and the SEA Community for helpful comments on the manuscript. We also thank the Western Carolina Spring 2024 bioinformatics class for their contributions to the Evcara genome annotation.

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)

- 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)
- Horton EN, Beach EK, Cook KT, Cronin KG, Haag AT, Salter SM, et al., Anders. 2024. Complete genome sequence of *Microbacterium foliorum* phage Curie, a podovirus isolated from soil in Spokane, Washington. *Microbiology Resource Announcements* 13: 10.1128/mra.00408-24. DOI: [10.1128/mra.00408-24](https://doi.org/10.1128/mra.00408-24)
- Jacobs-Sera D, Abad LA, Alvey RM, Anders KR, Aull HG, Bhalla SS, et al., Hatfull. 2020. Genomic diversity of bacteriophages infecting *Microbacterium* spp. *PLOS ONE* 15: e0234636. DOI: [10.1371/journal.pone.0234636](https://doi.org/10.1371/journal.pone.0234636)
- LaBianca K, Butura J, McKnight A, Pan K, Whelan S, Punska E, et al., Rocheleau. 2023. Complete genome sequence of phi29-like *Microbacterium foliorum* podovirus phage PineapplePizza. *Microbiology Resource Announcements* 12: 10.1128/mra.00478-23. DOI: [10.1128/mra.00478-23](https://doi.org/10.1128/mra.00478-23)
- 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)
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al., Rothberg. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376-380. DOI: [10.1038/nature03959](https://doi.org/10.1038/nature03959)
- 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. 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)
- 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)
- 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)
- Xu J, Wang D, Gui M, Xiang Y. 2019. Structural assembly of the tailed bacteriophage ϕ 29. *Nature Communications* 10: 10.1038/s41467-019-10272-3. DOI: [10.1038/s41467-019-10272-3](https://doi.org/10.1038/s41467-019-10272-3)
- Yehle CO. 1978. Genome-linked protein associated with the 5' termini of bacteriophage phi29 DNA. *Journal of Virology* 27: 776-783. DOI: [10.1128/jvi.27.3.776-783.1978](https://doi.org/10.1128/jvi.27.3.776-783.1978)
- 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: This work was supported by the Western Carolina University Department of Chemistry and Physics, and the SEA-PHAGES program.

Author Contributions: Hannah B. Bonogafsky: formal analysis, writing - original draft. Kelly M. Freer: formal analysis, writing - original draft. Taylor B. Sanderson: formal analysis, writing - original draft. Kira A. Yost: formal analysis, writing - original draft. Ayden J. Pyle: formal analysis, writing - review editing. Ginavieve Rowley: formal analysis. Maria D. Gainey: conceptualization, formal analysis, methodology, resources, supervision, validation, writing - original draft, writing - review editing, funding acquisition.

Reviewed By: Kristen Butela, Anonymous

History: Received October 25, 2024 **Revision Received** December 4, 2024 **Accepted** December 6, 2024 **Published Online** December 8, 2024 **Indexed** December 22, 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: Bonogafsky, HB; Freer, KM; Sanderson, TB; Yost, KA; Pyle, AJ; Rowley, G; Gainey, MD (2024). Complete annotated genome sequence of *Microbacterium paraoxydans* phage Evcara, a cluster GI podovirus isolated from compost.. *microPublication Biology*. [10.17912/micropub.biology.001398](https://doi.org/10.17912/micropub.biology.001398)

