

Complete Genome Sequences *Microbacterium paraoxydans* phages BarnCat, Pochacco, Smelly and SpiderBri and *Microbacterium foliorum* phage Crisis

Ana Acosta¹, Micah Baga¹, Kallysta Ballweber¹, Karenna Brand¹, Yehee Cho¹, Tanner Cox¹, Anheli Franco¹, Felipe Guevara Lopez¹, Felicity Hernandez¹, Lia Hutchins¹, Adam Johnson¹, Hyeri Kang¹, John Kang¹, Hojeong Kim¹, Mason Kim¹, Estrella Klar¹, Yongmin Kwon¹, Alexis Lee¹, Noah McMillan¹, Lauren Michel¹, Ruth Montiel¹, Shalom Mugenzi¹, Layla Murillo¹, Ethan Nguyen-Khoa¹, Traysen Nhem¹, Necker Peralta¹, Julianne Role¹, Natallie Ruiz¹, Annika Samayoa¹, Bridgette Sanchez¹, Ahmi Shin¹, Janelle Terriquez¹, Mateo Toledo¹, Zadie Tsao¹, Andrew Velasquez¹, Natasha Dean¹, Arturo Diaz^{1§}

¹Biology, La Sierra University, Riverside, California, United States

[§]To whom correspondence should be addressed: adiaz@lasierra.edu

Abstract

As part of the Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Sciences program, we report the discovery and genome sequence of microbacteriophages Crisis, BarnCat, Pochacco, Smelly, and SpiderBri. Phages were isolated from soil samples in Riverside, CA and Viroqua, WI using the hosts *Microbacterium foliorum* and *Microbacterium paraoxydans*.

Phage	Crisis	BarnCat	Pochacco	Smelly	SpiderBri
Isolation host	Microbacterium foliorum NRRL B-24224	Microbacterium paraoxydans NRRL B-14843	Microbacterium paraoxydans NRRL B-24275	Microbacterium paraoxydans NRRL B-24275	Microbacterium paraoxydans NRRL B-24275
Sample location (city, state)	Riverside, CA	Viroqua, WI	Riverside, CA	Riverside, CA	Riverside, CA
GPS coordinates	33.9116 N, 117.50198 W	43.494722 N, 90.863889 W	33.90076 N, 117.49401 W	33.90738 N, 117.50807 W	33.8946 N, 117.51979 W
Isolation method	Enriched	Enriched	Direct	Enriched	Enriched
Lysate Titer (pfu/ml)	4.9 x 10 ¹⁰	5 x 10 ⁹	5.4 x 10 ¹⁰	1.66 x 10 ¹⁰	2.5 x 10 ¹⁰
Approximate read coverage (fold, x)	90	638	10,921	10,907	4,143
# of 150 base single-end reads	31,873	268,38			
# of 100 base single-end reads			1,408,121	2,000,240	743,422
Genome length (bp)	52,987	61,511	17,880	17,859	17,868
GC content %	68.90%	58.10%	52.70%	52.90%	52.30%
Character of genome ends	Circularly permuted	Direct Terminal Repeat	Covalent Terminal Protein	Covalent Terminal Protein	Covalent Terminal Protein
# of ORFs (# with predicted function)	90 (29)	125 (32)	25 (12)	24 (12)	25 (12)
# of tRNAs	0	1	0	0	0
Capsid size range	71 nm (n=1)	65 nm (n=1)	32-37 nm (n=8)	50-53 nm (n=12)	37-38 nm (n=8)
Tail length range	135 nm (n=1)	174 nm (n=1)	9-10 nm (n=6)	15-17 nm (n=12)	10-11 nm (n=2)
Plaque size range	1-1.6 mm (n=10)	1-3 mm (n=50)	1-1.8 mm (n=10)	0.5-1.2 mm (n=10)	1.4-2.1 mm (n=10)
Cluster Assignment	EC	GB	GK	GK	GK
GenBank accession #	PP750965	PP978804	PV915814	PV15830	PV915808
SRA accession #	SRX30257450	SRX30257449	SRX30257453	SRX30257454	SRX30257455

Table 1. Phage Genome Information:

Genome characteristics of phages isolated in this study, including host, lysate titers, GC content, genome length, character of genome ends, number of predicted open reading frames (ORFs), number of predicted tRNAs, and phage particle characteristics. Phages were assigned to clusters based on gene content similarity of at least 35% to phages in the Actinobacteriophage database. Sequencing information for each phage as well as sample collection location and the phage isolation method are also listed.

Description

Comparative analysis of the genomes of phages has provided important insights into bacteriophage evolution and diversity. To expand the catalog of bacteriophages capable of infecting *Microbacterium*, five phages were isolated from soil samples collected in Southern California and Wisconsin using *Microbacterium foliorum* NRRL B-24224, and two *Microbacterium paraoxydans* species, NRRL B-14843 and NRRL B-24275, as hosts (Table 1) (Jacobs-Sera et al., 2020). Members of the genus *Microbacterium* are Gram-positive and commonly found in diverse environments, including soil, aquatic habitats, dairy products, and plant surfaces (Russell et al. 2019). Certain species, such as *M. foliorum* and *M. paraoxydans*, have also been isolated from clinical sources, including blood specimens and wound swabs (Laffineur et al. 2003; Gneiding et al. 2008).

Soil samples were initially suspended in peptone-yeast extract-calcium (PYCa) liquid medium and incubated at 30° C with agitation for 90 minutes. Following incubation, the mixtures were centrifuged, and the supernatants were filtered using a 0.22 µm membrane (Zorawik et al., 2024). For phages isolated via the "direct" method, 500 µl of each filtrate was mixed with 250 µl of *M. paraoxydans* in PYCa soft agar and incubated at 30° C for up to 48 hours to allow plaque development. In the "enrichment" method, filtrates were inoculated with either *M. paraoxydans* or *M. foliorum* and incubated with shaking at 30° C for 72 hours. These cultures, potentially enriched for microbacteriophages, were subsequently refiltered and plated in soft agar containing the corresponding bacterial host. Clear plaques were observed in all cases, and phages were purified by performing three rounds of plaque picking and plating. For each round of purification, plaques were picked that were at least 1 cm from the nearest plaque neighbor. High-titer lysates were prepared and used for negative-staining transmission electron microscopy (TEM) using uranyl acetate (Zorawik et al., 2024). Based on TEM imaging, phages Crisis and BarnCat exhibited siphoviral morphology, whereas Pochacco, Smelly, and SpiderBri displayed podoviral morphology. Phage lysate titers, capsid width, tail length, and plaque size ranges are listed in Table 1.

The phenol-chloroform DNA extraction protocol was optimized for high-titer lysate samples. Lysates were treated with RNase/DNase and Proteinase K, followed by incubation and phase separation using phenol:chloroform:isoamyl alcohol. DNA was precipitated with sodium chloride and ethanol, washed with 70% ethanol, air-dried, and resuspended in water. DNA quality and concentration were assessed using a NanoDrop spectrophotometer. Library preparation was performed using the NEB Ultra II FS Kit. For phages Crisis and BarnCat, sequencing was conducted on an Illumina MiSeq 1000 (v3 reagents), generating 150 bp single-end reads. For phages Pochacco, Smelly, and SpiderBri, sequencing was conducted on an Illumina NexSeq 1000 (XLEAP-P1 kit), generating 100 base reads. MiSeq raw reads were assembled without trimming whereas the NextSeq raw 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 (Jiang et al., 2014; Martin, 2011). Assembly of raw reads was carried out using Newbler v2.9 with default settings (Russell 2018), and completeness of the assembled genomes was manually assessed using Consed v29 (Gordon and Green 2013). Read counts and estimated genome coverage are detailed in Table 1. The resulting genome lengths ranged from 17,859 to 61,511 bp, with GC contents between 52.3% and 68.9% (Table 1).

Genome annotation was performed using the Phage Evidence Collection and Annotation Network (PECAAN) platform (Rinehart et al., 2016). Gene start and stop sites were predicted using a combination of Glimmer v3.02 (Delcher et al. 2007), Genemark v2.5 (Besemer and Borodovsky 2005), and Starterator v.546f (Pope & Jacobs-Sera, 2018). Functional predictions of encoded proteins were determined using a combination of HHpred (PDB_mmCIF70, UniProt, Pfam-A v.36, and NCBI Conserved Domain databases) (Söding et al. 2005), BLASTp v.2.16.0 (Altschul et al. 1997) alignments against the Actinobacteriophage protein (<https://phagesdb.org/>) (Russell & Hatfull, 2017) and NCBI non-redundant protein sequences databases (<https://blast.ncbi.nlm.nih.gov>), and Phamerator (Cresawn et al. 2011). Transmembrane helices were identified using DeepTMHMM v1.0.44 (Hallgren et al. 2022), TOPCONS v2.0 (Tsirigos et al. 2015), and SOSUI v1.11 (Hirokawa et al. 1998). tRNAs were detected using ARAGON v1.2.41 and tRNAscan-SE v2.0.12 (Laslett and Canback 2004; Lowe and Chan 2016). All software were used with default parameters.

No integrase or repressor functions could be identified across any of the genomes of these five phages, suggesting that these phages do not undergo lysogeny and are thus predicted to be lytic phages. Phages were assigned to clusters using the phagesDB Gene Content Similarity (GCS) tool, based on a minimum of 35% shared gene content with existing entries in the Actinobacteriophage database (Russell & Hatfull, 2017). Pochacco, Smelly, and SpiderBri are the inaugural members of cluster GK, and the genomes are characterized by the presence of a gene encoding a predicted terminal protein, which functions by covalently attaching to the phage genome termini where it can prime the initiation of DNA replication. Another notable feature shared by all three phages is that the predicted terminase gene is in the middle of the lysis cassette, downstream of two genes encoding lysin-associated domains, specifically an endolysin with a protease M15 domain followed by an endolysin with a protease M23 domain, and upstream of two transmembrane proteins. Additionally, all three have a relatively low GC content compared to most *Microbacterium* species (~52% vs ~67%) (Jacobs-Sera et al., 2020). Crisis was assigned to cluster EC and the genome is circularly permuted. As is the case for all cluster EC phages, the putative open reading frames in Crisis's genome are transcribed unidirectionally, and contain an asymmetric 14-bp sequence (CTATAGGTGTAAGC) upstream of transcriptional start sites (Barruga et al., 2024; Jacobs-Sera et al., 2020; Kim et al., 2022; Lin et al., 2020; Miller et al., 2019). BarnCat was assigned to cluster GB and the end of

the genome has direct terminal repeats. BarnCat encodes two putative endonuclease VII proteins, and four HNH endonucleases scattered throughout the genome. In comparison, the cluster GB phages Lifes and WaterT are not predicted to encode any HNH endonucleases, whereas Cassita and LeeroyJenkins encode one and two, respectively. Among the cluster GB phages, LeeroyJenkins is the only other phage predicted to encode a putative endonuclease VII.

Data Availability

GenBank and Sequence Read Archive (SRA) accession numbers are provided in Table 1.

Acknowledgements: We would like to thank Daniel Russell and Rebecca Garlena for assembling the genome, the SEA-PHAGES program for providing support on genome annotation and lysate archiving, and Ilkeun Lee at the Nanofabrication and Electron Microscopy Facility at UC Riverside for imaging the phage particles.

References

- Altschul S. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25: 3389-3402. DOI: [10.1093/nar/25.17.3389](https://doi.org/10.1093/nar/25.17.3389)
- Barruga B, Benitez S, Bouit E, Bouit T, Cagang E, Cantos J, et al., Diaz A. 2024. Complete Genome Sequences of Phages EarickHC, Figueroism, FinalFrontier, SBlackberry, Skylord, and Slay isolated using *Microbacterium foliorum*. *MicroPubl Biol* 2024: 10.17912/micropub.biology.001357. PubMed ID: [39712932](https://pubmed.ncbi.nlm.nih.gov/39712932/)
- 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)
- Gneiding K, Frodl R, Funke G. 2008. Identities of *Microbacterium* spp. Encountered in Human Clinical Specimens. *Journal of Clinical Microbiology* 46: 3646-3652. DOI: [10.1128/jcm.01202-08](https://doi.org/10.1128/jcm.01202-08)
- Gordon D, Green P. 2013. *Consed*: a graphical editor for next-generation sequencing. *Bioinformatics* 29: 2936-2937. DOI: [10.1093/bioinformatics/btt515](https://doi.org/10.1093/bioinformatics/btt515)
- Hallgren J, Tsirigos 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)
- 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)
- 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. DOI: [10.1186/1471-2105-15-182](https://doi.org/10.1186/1471-2105-15-182)
- Kim U, Paul ES, Diaz A. 2022. Characterization of Phages YuuY, KaiHaiDragon, and OneinaGillian Isolated from *Microbacterium foliorum*. *International Journal of Molecular Sciences* 23: 6609. DOI: [10.3390/ijms23126609](https://doi.org/10.3390/ijms23126609)
- Laffineur K, Avesani Vr, Cornu G, Charlier J, Janssens ML, Wauters G, Delmée M. 2003. Bacteremia Due to a Novel *Microbacterium* Species in a Patient with Leukemia and Description of *Microbacterium paraoxydans* sp. nov. *Journal of Clinical Microbiology* 41: 2242-2246. DOI: [10.1128/jcm.41.5.2242-2246.2003](https://doi.org/10.1128/jcm.41.5.2242-2246.2003)
- 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)
- Lin H, Reeves M, Acevedo M, Bass K, Chau E, Ching B, et al., Diaz. 2020. Genome Sequences of Bacteriophages ClearAsMud and Kauala, Isolated from *Microbacterium foliorum*. *Microbiology Resource Announcements* 9: 10.1128/mra.01026-20. DOI: [10.1128/mra.01026-20](https://doi.org/10.1128/mra.01026-20)
- 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)
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17: 10. DOI: [10.14806/ej.17.1.200](https://doi.org/10.14806/ej.17.1.200)

Miller G, Tran S, Abrahams R, Bazan D, Blaylock E, Choi J, et al., Diaz. 2019. Genome Sequences of Bacteriophages KaiHaiDragon and OneinaGillian, Isolated from *Microbacterium foliorum* in Riverside, California. *Microbiology Resource Announcements* 8: 10.1128/mra.00002-19. DOI: [10.1128/mra.00002-19](https://doi.org/10.1128/mra.00002-19)

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)

Rinehart, C. A., Gaffney, B., Wood, J., & Smith, J. (2016). *PECAAN, a phage evidence collection and annotation network*. <https://discover.kbrinsgd.org/login>

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, Garlena RA, Hatfull GF. 2019. Complete Genome Sequence of *Microbacterium foliorum* NRRL B-24224, a Host for Bacteriophage Discovery. *Microbiology Resource Announcements* 8: 10.1128/mra.01467-18. DOI: [10.1128/MRA.01467-18](https://doi.org/10.1128/MRA.01467-18)

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)

Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Research* 43: W401-W407. DOI: [10.1093/nar/gkv485](https://doi.org/10.1093/nar/gkv485)

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 La Sierra University College of Arts and Sciences and the Biology Department.

Author Contributions: Ana Acosta: investigation. Micah Baga: investigation. Kallysta Ballweber: investigation. Karenn Brand: investigation. Yehee Cho: investigation. Tanner Cox: investigation. Anhel Franco: investigation. Felipe Guevara Lopez: investigation. Felicity Hernandez: investigation. Lia Hutchins: investigation. Adam Johnson: investigation. Hyeri Kang: investigation. John Kang: investigation. Hojeong Kim: investigation. Mason Kim: investigation. Estrella Klar: investigation. Yongmin Kwon: investigation. Alexis Lee: investigation. Noah McMillan: investigation. Lauren Michel: investigation. Ruth Montiel: investigation. Shalom Mugenzi: investigation. Layla Murillo: investigation. Ethan Nguyen-Khoa: investigation. Traysen Nhem: investigation. Necker Peralta: investigation. Julianne Role: investigation. Natallie Ruiz: investigation. Annika Samayoa: investigation. Bridgette Sanchez: investigation. Ahmi Shin: investigation. Janelle Terriquez: investigation. Mateo Toledo: investigation. Zadie Tsao: investigation. Andrew Velasquez: investigation. Natasha Dean: supervision. Arturo Diaz: supervision, validation, writing - original draft, writing - review editing, investigation.

Reviewed By: Anonymous, Kristen Butela

History: Received September 4, 2025 **Revision Received** December 1, 2025 **Accepted** December 19, 2025 **Published Online** December 22, 2025 **Indexed** January 5, 2026

Copyright: © 2025 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: Acosta A, Baga M, Ballweber K, Brand K, Cho Y, Cox T, et al., Diaz A. 2025. Complete Genome Sequences *Microbacterium paraoxydans* phages BarnCat, Pochacco, Smelly and SpiderBri and *Microbacterium foliorum* phage Crisis. *microPublication Biology*. [10.17912/micropub.biology.001840](https://doi.org/10.17912/micropub.biology.001840)