





# Complete Genome Sequence of Finnry, a Subcluster L3 Mycobacteriophage from Charleston, South Carolina

 Christine A. Byrum,<sup>a</sup>  Véronique A. Delesalle,<sup>b</sup> Claudia L. Gold,<sup>a</sup> Daniel J. Bennett,<sup>a</sup> B. Conner Fox,<sup>a</sup> Brandon M. Houston,<sup>a</sup> Harrison E. Koller,<sup>a</sup> Peyton G. Russell,<sup>a</sup> Pavi Sreekumar,<sup>a</sup> Bradley R. Teasley,<sup>a</sup> Eva Vadoros,<sup>a</sup> Anastasia M. Zimmerman,<sup>a</sup> Mouna S. DiBenedetto,<sup>a</sup> Christopher A. Korey<sup>a</sup>

<sup>a</sup>Department of Biology, College of Charleston, Charleston, South Carolina, USA

<sup>b</sup>Department of Biological Sciences, Gettysburg College, Gettysburg, Pennsylvania, USA

**ABSTRACT** Subcluster L3 bacteriophage Finnry was isolated from soil collected in Charleston, South Carolina, using *Mycobacterium smegmatis* mc<sup>2</sup>155 as a host. The genome of this temperate siphovirus is 75,632 bp long (130 predicted protein-coding genes, 9 tRNAs, and no transfer-messenger RNAs), and BLASTn alignment revealed 99.86% identity with the genome of L3 mycobacteriophage Samty.

Undergraduates in the Howard Hughes Medical Institute (HHMI) Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program (1) studied the mycobacteriophage Finnry in a broader effort to characterize viral diversity/evolution and improve phage therapy approaches (2, 3). Finnry was obtained from dry, dusty soil at the College of Charleston, South Carolina (32.783445N, 79.937537W), and isolated in *Mycobacterium smegmatis* mc<sup>2</sup>155 using enrichment at 37°C followed by two purification/amplification cycles in 7H9 top agar, as described in the SEA-PHAGES Discovery Guide (4). Although Finnry forms clear plaques at 37°C, genome analysis indicated that the virus is temperate. Transmission electron microscopy revealed that the phage has *Siphoviridae* morphology, an icosahedral capsid, and a flexible, noncontractile tail (Fig. 1 and Table 1).

To extract genomic DNA from high-titer lysates, the Promega Wizard DNA cleanup system was used, and a DNA library was prepared with the NEBNext Ultra II DNA library prep kit. Pittsburgh Bacteriophage Institute sequenced Finnry on an Illumina MiSeq system (MiSeq reagent kit v3) (5), and 771,310 single-end reads (150 bp) were obtained. Raw reads were assembled into one contig with Newbler v2.9 (6) and verified with Consed v29.0 (7). Finnry's genome is 75,632 bp, with 1,492× coverage and a G+C content of 59.3%. Genome termini with 3' single-stranded extensions (5'-TCGATCAGCC) were identified using PAUSE (<https://cpt.tamu.edu/computer-resources/pause>).

Annotation was performed with the PECAAN (8) workflow tool, and final files were transferred to DNA Master v5.23.2 (<https://phagesdb.org/DNAMaster>). Programs utilized to identify putative genes included GLIMMER v3.02 (9), Phamerator Actino\_prophage v5 (10), GeneMark v3.25 (11), Starterator v1.1 (12), ARAGORN v1.2.38 (13), and tRNAscan-SE v3.0 (14). Functional assignments and domains were detected using BLASTp v2.8.1+ (15), HHpred (16), and the NCBI Conserved Domain Database (CDD) searched with reverse position specific (RPS)-BLAST from NCBI BLAST v2.8.1+ (17) (parameters at <https://seaphages.org/forums/topic/5398>). Default parameters were used for other software.

Finnry's genome contains 130 predicted protein-coding genes (51 with assigned putative functions), 9 tRNAs, and no transfer-messenger RNAs. Potential gene duplications include tandem duplication of the WhiB family transcription factor sequences gp79/gp80 (BLASTp indicated 37.66% identity and 79% query coverage) and displaced duplication of gp121/gp131 (BLASTp indicated 42.59% identity and 93% query coverage).

**Editor** Kenneth M. Stedman, Portland State University

**Copyright** © 2022 Byrum et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

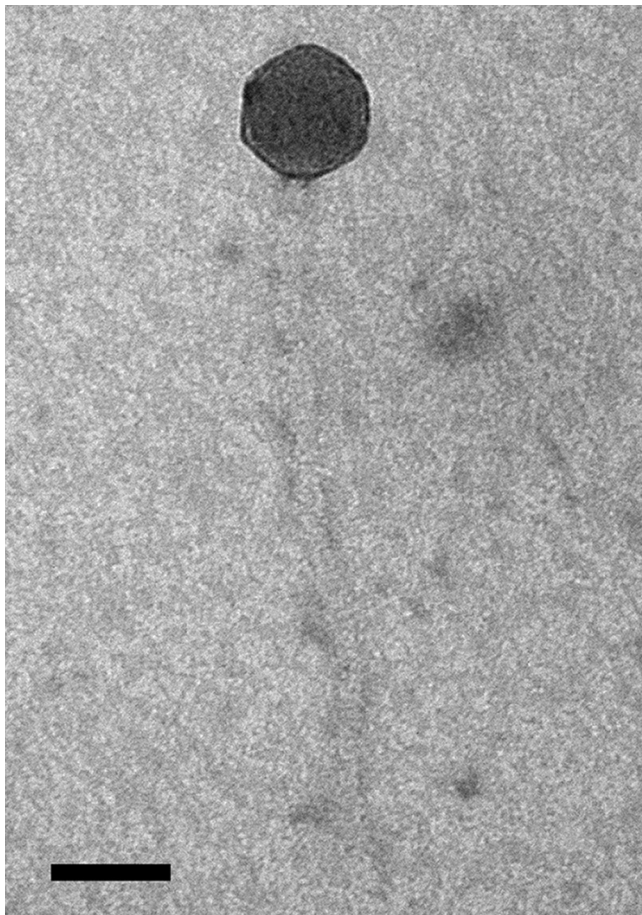
Address correspondence to Christine A. Byrum, [byrumc@cofc.edu](mailto:byrumc@cofc.edu).

The authors declare no conflict of interest.

**Received** 30 June 2022

**Accepted** 21 July 2022

**Published** 15 August 2022



**FIG 1** Morphology of the cluster L3 *Siphoviridae* member Finnry examined using a JEOL1010 transmission electron microscope (80 kV). High-titer lysates collected on Formvar-coated copper grids were negatively stained with 1% uranyl acetate (4). Scale bar, 50 nm.

Based on nucleotide similarity, Finnry is classified with similar phages into the L cluster/L3 subcluster, with cluster members sharing >50% nucleotide identity and/or >35% gene content similarity (GCS) (18–20). To compare the distribution of phamilies (phams) (potentially homologous protein-coding sequences sharing >32.5% amino acid identity in CLUSTALW and BLASTp E-values  $<10^{-50}$ ) between Finnry and related actinobacteriophages, Phamerator was used (10). Finnry's genome contains 9 phams unique to L3 subcluster members and also conserved in all L3 members (Table 1), 2 phams (gp134 and gp137) occurring in only one other L3 member, and 2 phams (gp130 and gp138) unique to Finnry.

GCS scores (19) and whole-genome BLASTn alignments (15) revealed that Finnry's genome is most similar to that of Samty (93.4% GCS, 99.86% identity, and 99% query coverage), an L3 bacteriophage from Huntsville, Texas. Most L3 subcluster phages (15/16 phages) occur in the southeastern United States (Florida, South Carolina, Louisiana, and Texas) (11 phages) or South Africa (4 phages). Whirlwind is from Pittsburgh, Pennsylvania.

**Data availability.** The GenBank and SRA accession numbers for Finnry are presented in Table 1.

#### ACKNOWLEDGMENTS

This study was generously supported by the HHMI SEA-PHAGES program, as well as the College of Charleston Department of Biology and SC INBRE (publication costs were covered by a SC INBRE program grant from the National Institutes of Health National Institute of General Medical Sciences [grant P20GM103499-20]).

**TABLE 1** Characteristics of the Finrry bacteriophage

Parameter	Data for Finrry
GenBank accession no.	<a href="#">MN096377</a>
SRA accession no.	<a href="#">SRX14989442</a>
Collection site	Charleston, South Carolina, USA
Collection site coordinates	32.783445N, 79.937537W
Isolation host	<i>Mycobacterium smegmatis</i> mc <sup>2</sup> 155
Genome size (bp)	75,632
Coverage (×)	1,492
G+C content (%)	59.3
No. of predicted protein-coding genes	130
No. of tRNAs	9
No. of transfer-messenger RNAs	0
Morphotype	<i>Siphoviridae</i>
Subcluster	L3
Plaque size (mm) ( <i>n</i> = 10)	
Range	0.67–1.10
Mean	0.92
Capsid size (nm) ( <i>n</i> = 3) <sup>b</sup>	
Range	54.5–56.8
Mean	55.7
Tail length (nm) ( <i>n</i> = 3) <sup>b</sup>	
Range	286.4–307.7
Mean	299.5
Tail width (nm) ( <i>n</i> = 3) <sup>b</sup>	
Range	13.6–14.3
Mean	13.83
Predicted Finrry protein-coding genes (phams) that are unique to and conserved in all L3 subcluster members (all with no known function) <sup>a</sup>	4, 35, 45, 47, 49, 58, 105, 106, 117

<sup>a</sup> Based on data available in Phamerator on 16 June 2022 (10).

<sup>b</sup> Measurements acquired from transmission electron micrographs.

We thank Graham F. Hatfull, Deborah Jacobs-Sera, Welkin H. Pope, Daniel A. Russell, and Rebecca A. Garlena at the University of Pittsburgh for sequencing, quality control, and assembly of the genome, as well as related training for the project, and Nancy Smythe at the Medical University of South Carolina Department of Pathology for assistance with the morphological characterization using electron microscopy.

## REFERENCES

- Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, Dennehy JJ, Denver DR, Dunbar D, Elgin SCR, Findley AM, Gissendanner CR, Golebiewska UP, Guild N, Hartzog GA, Grillo WH, Hollowell GP, Hughes LE, Johnson A, King RA, Lewis LO, Li W, Rosenzweig F, Rubin MR, Saha MS, Sandoz J, Shaffer CD, Taylor B, Temple L, Vazquez E, Ware VC, Barker LP, Bradley KW, Jacobs-Sera D, Pope WH, Russell DA, Cresawn SG, Lopatto D, Bailey CP, Hatfull GF. 2014. A broadly implementable research course in phage discovery and genomics for first-year undergraduate students. *mBio* 5:e01051-13. <https://doi.org/10.1128/mBio.01051-13>.
- Diacon AH, Guerrero-Bustamante CA, Rosenkranz B, Rubio Pomar FJ, Vanker N, Hatfull GF. 2022. Mycobacteriophages to treat tuberculosis: dream or delusion? *Respiration* 101:1–15. <https://doi.org/10.1159/000519870>.
- Nick JA, Dedrick RM, Gray AL, Vladar EK, Smith BE, Freeman KG, Malcolm KC, Epperson LE, Hasan NA, Hendrix J, Callahan K, Walton K, Vestal B, Wheeler E, Rysavy NM, Poch K, Caceres S, Lovell VK, Hisert KB, de Moura VC, Chatterjee D, De P, Weakly N, Martiniano SL, Lynch DA, Daley CL, Strong M, Jia F, Hatfull GF, Davidson RM. 2022. Host and pathogen response to bacteriophage engineered against *Mycobacterium abscessus* lung infection. *Cell* 185:1860–1874. <https://doi.org/10.1016/j.cell.2022.04.024>.
- Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull G. 2018. Phage discovery guide. Howard Hughes Medical Institute, Chevy Chase, MD. <https://seaphagesphagediscoveryguide.helpdocsonline.com/home>.
- Russell DA. 2018. Sequencing, assembling, and finishing complete bacteriophage genomes. *Methods Mol Biol* 1681:109–125. [https://doi.org/10.1007/978-1-4939-7343-9\\_9](https://doi.org/10.1007/978-1-4939-7343-9_9).
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen Y-J, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer MLI, Jarvie TP, Jirage KB, Kim J-B, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376–380. <https://doi.org/10.1038/nature03959>.
- Gordon D, Green P. 2013. Consed: a graphical editor for next-generation sequencing. *Bioinformatics* 29:2936–2937. <https://doi.org/10.1093/bioinformatics/btt515>.
- 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. [https://seaphages.org/media/docs/PECAAN\\_User\\_Guide\\_Dec7\\_2016.pdf](https://seaphages.org/media/docs/PECAAN_User_Guide_Dec7_2016.pdf).
- Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res* 27:4636–4641. <https://doi.org/10.1093/nar/27.23.4636>.
- 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. <https://doi.org/10.1186/1471-2105-12-395>.
11. Lukashin A, Borodovsky M. 1998. Genemark.hmm: new solutions for gene finding. *Nucleic Acids Res* 26:1107–1115. <https://doi.org/10.1093/nar/26.4.1107>.
  12. Pacey M. 2016. Pope W (ed). Starterator guide. University of Pittsburgh, Pittsburgh, PA. [https://seaphages.org/media/docs/Starterator\\_Guide\\_2016.pdf](https://seaphages.org/media/docs/Starterator_Guide_2016.pdf).
  13. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 32:11–16. <https://doi.org/10.1093/nar/gkh152>.
  14. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955–964. <https://doi.org/10.1093/nar/25.5.955>.
  15. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
  16. Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33:W244–W248. <https://doi.org/10.1093/nar/gki408>.
  17. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015. CDD: NCB's Conserved Domain Database. *Nucleic Acids Res* 43:D222–D226. <https://doi.org/10.1093/nar/gku1221>.
  18. Hatfull GF, Jacobs-Sera D, Lawrence JG, Pope WH, Russell DA, Ko C-C, Weber RJ, Patel MC, Germane KL, Edgar RH, Hoyte NN, Bowman CA, Tantoco AT, Paladin EC, Myers MS, Smith AL, Grace MS, Pham TT, O'Brien MB, Vogelsberger AM, Hryckowian AJ, Wynalek JL, Donis-Keller H, Bogel MW, Peebles CL, Cresawn SG, Hendrix RW. 2010. Comparative genomic analysis of 60 mycobacteriophage genomes: genome clustering, gene acquisition and gene size. *J Mol Biol* 397:119–143. <https://doi.org/10.1016/j.jmb.2010.01.011>.
  19. Mavrich TN, Hatfull GF. 2017. Bacteriophage evolution differs by host, life-style and genome. *Nat Microbiol* 2:17112. <https://doi.org/10.1038/nmicrobiol.2017.112>.
  20. Hatfull GF. 2020. Actinobacteriophages: genomics, dynamics, and applications. *Annu Rev Virol* 7:37–61. <https://doi.org/10.1146/annurev-virology-122019-070009>.