

Evaluation of Genome Sequences of the Bacteriophages JeTaime and Luna22

Maximiliano F. Flota,^a Caitlyn R. Moss,^a  Mitchell F. Balish,^b Nikki Chen,^a Kristen C. Cherry,^a Kathleen A. Cornely,^c Hayley M. D'Alessandro,^a Mouna S. DiBenedetto,^a Joseph A. DeGiorgis,^c Steven Grant Dixon,^a Emily G. Dombrowski,^a Megan K. Edwards,^a Jonathan C. Eskew-Martin,^a Isabel E. P. Finnegan,^a Abanob G. Hanna,^a Sarah E. Hunter,^a Sabrina L. Johnson,^a Virginia A. L. Kenan,^a Chanelle Kendrick,^c Lucas C. Licaj,^a Victoria C. Maldonado,^a Megan G. Mazzei,^a Gracen E. Mitrick,^a Bradford L. J. Nelson,^a Jui S. Patel,^a Alexander I. Parry,^a Kacey M. Smekrud,^a Kirsten K. Snyder,^a Jonathan A. Stewart,^a Fredrick K. Swiger,^a Madison K. Thomas,^a Josiah C. Waters,^a  Christine A. Byrum^a

^aDepartment of Biology, College of Charleston, Charleston, South Carolina, USA

^bDepartment of Microbiology, Miami University, Oxford, Ohio, USA

^cDepartment of Chemistry and Biochemistry, Providence College, Providence, Rhode Island, USA

Maximiliano F. Flota and Caitlyn R. Moss share equal credit for first authorship. Authorship order is alphabetical for these two individuals.

ABSTRACT The mycobacteriophages JeTaime (E cluster) and Luna22 (Q cluster) were isolated from soil in Providence, Rhode Island, and Charleston, South Carolina, respectively, using a *Mycobacterium smegmatis* mc² 155 host. The genome of JeTaime is 75,099 bp (142 predicted genes), and that of Luna22 is 53,730 bp (87 predicted genes). Both phages exhibit *Siphoviridae* morphology.

Bacteriophage genome characterization provides key insights into viral evolution and genetic diversity for both basic and applied research. The mycobacteriophages JeTaime and Luna22 were investigated in partnership with the Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program through the Howard Hughes Medical Institute (HHMI) (1). The phage JeTaime was found in peaty surface soil from Providence, Rhode Island, and Luna22 was found in loamy surface soil from Charleston, South Carolina (Table 1). Both phages were isolated using enrichment at 37°C, followed by purification and amplification in *Mycobacterium smegmatis* mc² 155, as described in the SEA-PHAGES Phage Discovery Guide (2). Cultured viruses formed clear, circular plaques, and electron microscopy revealed *Siphoviridae* morphotypes with icosahedral capsids and long flexible tails (Fig. 1).

The Promega Wizard DNA cleanup system was used to extract DNA from phage lysates, and sequencing libraries were prepared with a NEBNext Ultra II DNA library prep kit (v3). Libraries for each phage were independently generated, and genomes were sequenced separately. Sequencing was performed at the Pittsburgh Bacteriophage Institute on an Illumina MiSeq system (MiSeq reagent kit v3) (3), and 150-bp single-end reads were collected (73,757 reads for JeTaime and 878,986 reads for Luna22). Raw reads were then assembled *de novo* into a single contig using Newbler v2.9 (4) and verified with Consed v29.0 (5), as described by Russell (3). Using PAUSE (<https://cpt.tamu.edu/computer-resources/pause>), genome termini with 3' cohesive overhangs were identified, as follows: JeTaime, 5'-CGTTGTCA-3'; Luna22, 5'-TAAGCCGCGGGTA-3'. Table 1 summarizes additional findings.

Each genome was annotated on PECAAN (6) using GLIMMER v3.02 (7), GeneMark v2.5 (8), and Starterator v1.1 (9). To refine start sites, to determine clusters, and to compare related genomes, Phamerator (10, 11) was utilized. Functional assignments were made with BLASTp v2.9 (12), HHpred (13), TMHMM2 (<http://www.cbs.dtu.dk/services/TMHMM>), TOPCONS v2 (14), and the NCBI Conserved Domain Database (CDD) (15). tRNAs and transfer-messenger RNAs were identified with ARAGORN v1.2.38 (16) and tRNAscan-SE v3.0 (17), and completed

Citation Flota MF, Moss CR, Balish MF, Chen N, Cherry KC, Cornely KA, D'Alessandro HM, DiBenedetto MS, DeGiorgis JA, Dixon SG, Dombrowski EG, Edwards MK, Eskew-Martin JC, Finnegan IEP, Hanna AG, Hunter SE, Johnson SL, Kenan VAL, Kendrick C, Licaj LC, Maldonado VC, Mazzei MG, Mitrick GE, Nelson BLJ, Patel JS, Parry AI, Smekrud KM, Snyder KK, Stewart JA, Swiger FK, Thomas MK, Waters JC, Byrum CA. 2021. Evaluation of genome sequences of the bacteriophages JeTaime and Luna22. *Microbiol Resour Announc* 10:e00746-21. <https://doi.org/10.1128/MRA.00746-21>.

Editor Kenneth M. Stedman, Portland State University

Copyright © 2021 Flota 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.

Received 22 July 2021

Accepted 2 September 2021

Published 7 October 2021

TABLE 1 Characteristics of the JeTaime and Luna22 phages

Phage name	GenBank accession no. (SRA no.)	Location	Genome size (bp)	Coverage (x)	% GC	No. of tRNAs (no. of tmRNAs)	Avg plaque diameter (mm)	Capsid size (nm)	Tail length (nm)
JeTaime	MZ322024.1 (SRX11158995)	Providence, RI (41.840000N, 71.40000W)	75,099	1,389	62.9	2 (0)	1.5–2.9, $\bar{x}=1.9$ (n = 10)	75.0–78.0, $\bar{x}=75.4$ (n = 4)	275–312.5, $\bar{x}=296$ (n = 3)
Luna22	MW924639 (SRX11158996)	Charleston, SC (32.784094N, 79.935882W)	53,730	2,352	65.7	0 (0)	2.0–3.0, $\bar{x}=2.5$ (n = 7)	62.5–75.0, $\bar{x}=70$ (n = 5)	231.3–262.5, $\bar{x}=244$ (n = 5)

files were transferred to DNA Master v5.22.2 (<https://phagesdb.org/DNAMaster>). All programs used default settings.

JeTaime is an E-cluster phage with 142 predicted protein-coding genes (54 assigned functions), while Luna22 is a Q-cluster phage with 87 predicted protein-coding genes (42 assigned functions) (11). Interestingly, sequence upstream of the single-stranded DNA (ssDNA)-binding protein (gp99) is absent in the JeTaime and NelitzamV (GenBank accession number [KT222941](#)) genomes. The other 103 published E-cluster members typically contain endonuclease VII and/or DNA methylase in that genome region. Also, the JeTaime gp89 (unidentified function) is rare, occurring in only one other E-cluster bacteriophage, 244 (GenBank accession number [DQ398041](#)). Whole-genome BLASTn alignment (12) revealed that the JeTaime genome is most similar to genomes of the E-cluster phages Cjw1 (GenBank accession number [NC_004681.1](#)) (99.24% identity and 96% coverage), Phrux (GenBank accession number [NC_021309.1](#)) (99.20% identity and 95% coverage), and ShereKhan (GenBank accession number [MH513983.1](#)) (99.10% identity and 95% coverage).

Interestingly, Luna22 and other Q-cluster genomes exhibit very high levels of conservation. Whole-genome BLASTn alignment (12) revealed that 13/15 published Q-cluster genomes share >99.80% identity with the Luna22 genome (100% coverage). Q-cluster phages from as far away as Washington state (Ein37) (GenBank accession number [MT114159](#)) and New Zealand (Daegal) (GenBank accession number [MK494095](#)) retain >98% nucleotide sequence identity to the Luna22 genome. Also, the genome of Giles (GenBank accession number [EU203571](#)) is nearly identical to that of Luna22 (99.96% identity and 100% coverage) (18–21); the only difference is that bases 33464 to 33479 in Giles are missing in Luna22. Loss of this sequence results in a frameshift mutation, causing gp42 (DNA-binding protein) to be truncated in Luna22 after amino acid 215.

Data availability. Individual GenBank and SRA numbers are listed in Table 1.

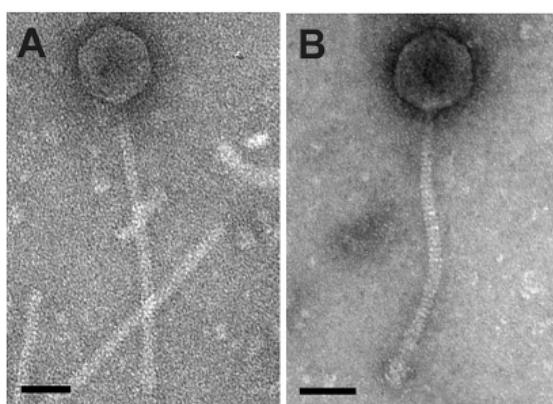


FIG 1 JeTaime (A) and Luna22 (B) morphology, examined using transmission electron microscopy. High-titer lysates placed on Formvar-coated grids were negatively stained with 1% uranyl acetate (2). JeTaime was imaged using a JEOL 20 \times transmission electron microscope (100 kV), and Luna22 was imaged using a JEOL 1010 transmission electron microscope (80 kV). Bars, 50 nm.

ACKNOWLEDGMENTS

This study was generously supported by the HHMI SEA-PHAGES program, the College of Charleston Department of Biology, the Providence College School of Arts and Sciences, the National Science Foundation under EPSCoR Cooperative Agreement OIA-1655221, and the National Center for Research Resources (grant 5 P20 RR016461) and the National Institute of General Medical Sciences (grant 8 P20 GM103499) of the National Institutes of Health.

We thank Nancy Smythe (Medical University of South Carolina) for imaging the Luna22 virus and Madison Dinkins for assistance with the annotation, as well as Graham F. Hatfull, Deborah Jacobs-Sera, Daniel A. Russell, Welkin H. Pope, and Rebecca A. Garlena at the University of Pittsburgh for sequencing, quality control, and assembly of the genome.

REFERENCES

1. 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>.
2. Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull G. 2018. Phage discovery guide. Howard Hughes Medical Institute, Chevy Chase, MD. <http://seaphagesphagediscoveryguide.helpdocsonline.com/home>.
3. 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.
4. 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>.
5. Gordon D, Green P. 2013. Consed: a graphical editor for next-generation sequencing. *Bioinformatics* 29:2936–2937. <https://doi.org/10.1093/bioinformatics/btt515>.
6. 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.
7. 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>.
8. Lukashin AV, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res* 26:1107–1115. <https://doi.org/10.1093/har/26.4.1107>.
9. Pacey M. 2016. Starterator guide. University of Pittsburgh, Pittsburgh, PA. https://seaphages.org/media/docs/Starterator_Guide_2016.pdf.
10. 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. Mavrich TN, Hatfull GF. 2017. Bacteriophage evolution differs by host, lifestyle, and genome. *Nat Microbiol* 2:17112. <https://doi.org/10.1038/nmicrobiol.2017.112>.
12. 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).
13. 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>.
14. 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 Res* 43:W401–W407. <https://doi.org/10.1093/nar/gkv485>.
15. 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: NCBI's Conserved Domain Database. *Nucleic Acids Res* 43:D222–D226. <https://doi.org/10.1093/nar/gku1221>.
16. 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>.
17. 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>.
18. Morris P, Marinelli LJ, Jacobs-Sera D, Hendrix RW, Hatfull GF. 2008. Genomic characterization of mycobacteriophage Giles: evidence for phage acquisition of host DNA by illegitimate recombination. *J Bacteriol* 190:2172–2182. <https://doi.org/10.1128/JB.01657-07>.
19. Dedrick RM, Marinelli LJ, Newton GL, Pogliano K, Pogliano J, Hatfull GF. 2013. Functional requirements for bacteriophage growth: gene essentiality and expression in mycobacteriophage Giles. *Mol Microbiol* 88:577–589. <https://doi.org/10.1111/mmi.12210>.
20. Mehla J, Dedrick RM, Caufield JH, Siefring R, Mair M, Johnson A, Hatfull GF, Uetz P. 2015. The protein interactome of mycobacteriophage Giles predicts functions for unknown proteins. *J Bacteriol* 197:2508–2516. <https://doi.org/10.1128/JB.00164-15>.
21. Mehla J, Dedrick RM, Caufield JH, Wagemans J, Sakhawalkar N, Johnson A, Hatfull GF, Uetz P. 2017. Virus-host protein-protein interactions of mycobacteriophage Giles. *Sci Rep* 7:16514. <https://doi.org/10.1038/s41598-017-16303-7>.