



# Complete Genome Sequences of *Mycobacterium smegmatis* Phages NihilNomen and Carlyle, Isolated in Las Vegas, Nevada

Alicia Salisbury,<sup>a</sup> Ryan Doss,<sup>a</sup> Astha Mehta,<sup>a</sup> Khadija Bhatti,<sup>a</sup> Ciera Dapra,<sup>a</sup> Audrey Huntsinger,<sup>a</sup> Stephanie Rodriguez,<sup>a</sup> Scott Yacek,<sup>a</sup> Rylee Sandberg,<sup>a</sup> Alexis Gildore,<sup>a</sup> Jacinda Knudtson,<sup>a</sup> Frances Tibayan,<sup>a</sup> Tiannah Ohta,<sup>a</sup> Neha Zafar,<sup>a</sup> Guadalupe Mercado,<sup>a</sup> Alan Le,<sup>a</sup> Natalie Mekhaeel,<sup>a</sup> Justin Willer,<sup>a</sup> Edith Rodrich-Zuniga,<sup>a</sup> Merissa McFarland,<sup>a</sup> Kurt Regner,<sup>a</sup> Christy Strong,<sup>a</sup> Philippos K. Tsourkas<sup>a</sup>

<sup>a</sup>School of Life Sciences, University of Nevada Las Vegas, Las Vegas, Nevada, USA

**ABSTRACT** We present the complete genomes of the *Mycobacterium smegmatis* phages Carlyle and NihilNomen, isolated from soil in Las Vegas, Nevada. The phages were isolated and annotated by undergraduate students enrolled in the Phage Discovery course offered by the School of Life Sciences at the University of Nevada Las Vegas.

Phages that infect *Mycobacterium smegmatis* account for the largest number of sequenced phage genomes, numbering approximately 1,800 (1). This is largely due to the versatility of this host and its consequent popularity with the Howard Hughes Medical Institute's (HHMI) Science Education Alliance-Phage Hunters Advancing Genomic and Evolutionary Science (SEA-PHAGES) program (2). Recently, three *M. smegmatis* phages isolated by students in the SEA-PHAGES program were used to treat a potentially lethal infection of antibiotic-resistant *Mycobacterium abscessus* in a cystic fibrosis patient (3). Here, we present the complete genomes of two *M. smegmatis* phages isolated by students enrolled in the Phage Discovery course (BIOL 207X and BIOL 209X) at the University of Nevada Las Vegas (UNLV).

Both phages were isolated from soil obtained from the community gardens on the UNLV campus by students in the course BIOL 207X. A direct isolation method was used. Environmental samples were incubated with enrichment broth and shaken (250 rpm, 2 h) at room temperature, followed by centrifugation and filter sterilization (0.22  $\mu\text{m}$ ) of the supernatant per HHMI's SEA-PHAGES Phage Discovery Guide (<https://seaphages.org/faculty/information/#phagediscovery>). The phages were then purified and amplified in *M. smegmatis* mc<sup>2</sup>155. The DNA was extracted using the manufacturer's protocol, provided in the phage DNA isolation kit (catalog number 46800; Norgen Biotek). Phage genomes were sequenced at the University of Pittsburgh. Sequencing libraries were prepared from genomic DNA by using an NEB Ultra II kit, producing 150-bp single-end reads. The libraries were sequenced with an Illumina MiSeq instrument, yielding single-end reads sufficient to provide at least 150-fold coverage for each genome. The reads were quality trimmed and assembled *de novo* by using Newbler version 2.9 with default settings, in each case yielding a single contig, which was checked for completeness, accuracy, and phage genomic termini by using Consed version 29, as described by Russell (4).

The assembly results (coverage depth, genome length, GC content, number of genes) and the phages' GenBank and SRA accession numbers are shown in Table 1. The phages were assigned to clusters based on genomic sequence similarity using the [phagesdb.org](http://phagesdb.org) database and Phamerator software with default settings (1, 5). Despite their similar geographic provenance, the phages are not closely related. Carlyle is cluster A1 and NihilNomen is cluster J based on the average nucleotide sequence identity. A ClustalW multiple alignment using default settings showed that the phages

**Citation** Salisbury A, Doss R, Mehta A, Bhatti K, Dapra C, Huntsinger A, Rodriguez S, Yacek S, Sandberg R, Gildore A, Knudtson J, Tibayan F, Ohta T, Zafar N, Mercado G, Le A, Mekhaeel N, Willer J, Rodrich-Zuniga E, McFarland M, Regner K, Strong C, Tsourkas PK. 2019.

Complete genome sequences of *Mycobacterium smegmatis* phages NihilNomen and Carlyle, isolated in Las Vegas, Nevada. Microbiol Resour Announc 8:e00677-19. <https://doi.org/10.1128/MRA.00677-19>.

**Editor** Simon Roux, DOE Joint Genome Institute

**Copyright** © 2019 Salisbury 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 Philippos K. Tsourkas, [philippos.tsourkas@unlv.edu](mailto:philippos.tsourkas@unlv.edu).

**Received** 6 August 2019

**Accepted** 26 August 2019

**Published** 19 September 2019

**TABLE 1** Phage GenBank and SRA accession numbers and genome assembly results

Phage name	GenBank accession no.	SRA accession no.	Avg coverage	Cluster	Genome length (bp)	GC content (%)	No. of genes
Carlyle	<a href="#">MK967401</a>	<a href="#">SRX5960826</a>	1,929	A1	51,220	63.6	91
NihilNomen	<a href="#">MK967402</a>	<a href="#">SRX5960825</a>	897	J	11,0439	60.8	241

have roughly 45% average nucleotide sequence identity (ANI) between them, which is the same as that of two randomly generated DNA sequences. Both phages are predicted to use the “cohesive ends with 3′ overhangs” DNA packaging strategy (6).

The assembled genomes were annotated using DNA Master with default settings, as described by Pope and Jacobs-Sera (7), by students in the course BIOL 209X at UNLV. We identified 91 genes in Carlyle and 241 in NihilNomen, of which 1 is a tRNA (Gly-tcc). The protein functions were assigned using Protein BLAST ([blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins](#)), Batch Web CD-Search ([ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi](#)), and HHpred ([toolkit.tuebingen.mpg.de/tools/hhpred](#)) using default settings. Using a cutoff E value of 1E-7, we assigned putative functions to 46 genes in Carlyle (50.5%) and 65 genes in NihilNomen (27%). A small and large terminase, a portal protein, a capsid maturation protease, a major capsid protein, a major tail protein, two tail assembly chaperones, a tail tape measure protein, lysin A and lysin B, and integrase were identified in both phages. The tail assembly chaperones have a predicted translational frameshift located in the 3′ region of the upstream tail assembly protein. A holin, excise, and Cro were identified in NihilNomen but not in Carlyle. Of interest is the identification of a 690-bp-long third terminase (gp2) in NihilNomen, an unusual feature present in some cluster J *M. smegmatis* phages, which is currently being investigated in our laboratory.

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

## ACKNOWLEDGMENTS

K.R., C.S., and P.K.T. acknowledge support from the Howard Hughes Medical Institute, the University of Pittsburgh (in particular, Becky Garlena, Dan Russell, Graham Hatfull, and Welkin Pope), and the School of Life Sciences and College of Science at the University of Nevada, Las Vegas. K.R. acknowledges support from the National Institute of General Medical Sciences (GM103440; NV INBRE).

K.R., C.S., and P.K.T. also acknowledge Brigham Young University, Julianne Grose, and Daniel Thompson for help in obtaining scanning electron microscopy (SEM) micrographs.

## REFERENCES

- Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. *Bioinformatics* 33:784–786. <https://doi.org/10.1093/bioinformatics/btw711>.
- 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>.
- Dedrick RM, Guerrero-Bustamante CA, Garlena RA, Russell DA, Ford K, Harris K, Gilmour KC, Soothill J, Jacobs-Sera D, Schooley RT, Hatfull GF, Spencer H. 2019. Engineered bacteriophages for treatment of a patient with a disseminated drug resistant *Mycobacterium abscessus*. *Nat Med* 25:730–733. <https://doi.org/10.1038/s41591-019-0437-z>.
- 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).
- Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF. 2011. Phamerator: a bioinformatic tool for comparative bacteriophage genomics. *BMC Bioinformatics* 12:395–409. <https://doi.org/10.1186/1471-2105-12-395>.
- Casjens SR, Gilcrease EB. 2009. Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions. *Methods Mol Biol* 502:91–111. [https://doi.org/10.1007/978-1-60327-565-1\\_7](https://doi.org/10.1007/978-1-60327-565-1_7).
- Pope WH, Jacobs-Sera D. 2018. Annotation of bacteriophage genome sequences using DNA Master: an overview. *Methods Mol Biol* 1681: 217–229. [https://doi.org/10.1007/978-1-4939-7343-9\\_16](https://doi.org/10.1007/978-1-4939-7343-9_16).