AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Isolation and Annotation of the Genome Sequences of Bacteriophages InvictusManeo (Subcluster K5) and Netyap (Subcluster L2)

Maximiliano F. Flota,[a] Véronique A. Delesalle,[b] Caitlyn R. Moss,[a] Davis S. Beeson,[a] Andrew M. Bogatkevich,[a] Clarissa C. Burgess,[a] Noemi Carretero Lazcano,[a] Jayda A. Carroll-Deaton,[a] Hayden A. Deprill,[a] Ivy S. Dickens,[a] Maria D. Gainey,[c] Sophia G. Gierszal,[a] Avery A. Goff,[a] Brooke K. Harris,[a] John B. LeBrun,[a] Jim Lin,[a] Molly R. McLaughlin,[a] Brian C. Metts,[a] Kristen L. O'Rear,[a] Maria A. Osorio Hernandez,[a] Isabella E. Raieta,[a] Erica D. Schmidt,[a] Thomas D. Sinkway,[a] Kimberly S. Sok,[a] Michael A. Ulrich,[a] Isaiah T. Velez,[a] Jamie R. Wallen,[c] Ayden R. Wardius,[a] Christine A. Byrum[a]

[a]Department of Biology, College of Charleston, Charleston, South Carolina, USA
[b]Department of Biological Sciences, Gettysburg College, Gettysburg, Pennsylvania, USA
[c]Department of Chemistry and Physics, Western Carolina University, Cullowhee, North Carolina, USA

ABSTRACT The mycobacteriophages InvictusManeo (K5 subcluster) and Netyap (L2 subcluster) were isolated from soils in Cullowhee Creek, Cullowhee, North Carolina. Both exhibit *Siphoviridae* morphology and infect *Mycobacterium smegmatis* mc²155. The InvictusManeo genome is 61,147 bp and contains 96 predicted protein-coding genes, whereas the Netyap genome is 76,366 bp with 131 predicted protein-coding genes.

The mycobacteriophages InvictusManeo and Netyap were isolated from single soil samples collected near Cullowhee Creek in Cullowhee, North Carolina (Table 1). Their genomes were analyzed in the Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) Program of the Howard Hughes Medical Institute (HHMI) (1) to increase understanding of viral evolution and diversity. Both viruses infect *Mycobacterium smegmatis* mc²155 and were isolated using enrichment at 37°C followed by two purification/amplification cycles in 7H9 top agar (SEA-PHAGES Phage Discovery Guide protocol) (2). Electron microscopy revealed that they exhibit *Siphoviridae* morphotypes (Fig. 1).

For sequencing, DNA was extracted from high-titer lysates with the Promega Wizard DNA cleanup system, followed by library preparation with a NEBNext Ultra II DNA library prep kit. The Western Carolina University Biotechnology Core performed shotgun sequencing on an Illumina MiSeq system (Nano v2 reagents) (3), producing 224,694 (InvictusManeo) and 238,646 (Netyap) single-end 150-bp reads. Reads were assembled into single contigs with Newbler v2.9 (4) and verified using Consed v29.0 (5) as described by Russell (3). Both genomes are linear with 3′ sticky overhangs (InvictusManeo, 5′-CTCAGTGGCAT-3′; Netyap, 5′-TCGATCAGCC-3′) and were annotated using the PECAAN workflow tool (6) and then transferred to DNA Master v5.22.2 (https://phagesdb.org/DNAMaster). GeneMark v2.5 (7), GLIMMER v3.02 (8), and Starterator v1.1 (9) were utilized to refine start sites, and comparative analysis was performed using Phamerator (10). Functional assignments were made with BLASTp v2.9 (11), HHpred (12), TMHMM v2.0 (https://services.healthtech.dtu.dk/service.php?TMHMM-2.0), TOPCONS v2 (13), and the NCBI Conserved Domain Database (CDD) (14), while tRNAs and transfer-messenger RNAs (tmRNAs) were identified using ARAGORN v1.2.38 (15) and tRNAscan-SE v3.0 (16). All programs used default parameters.

Mycobacteriophages sharing >50% nucleotide sequence similarity are categorized as members of the same cluster and are divided into subclusters based on average nucleotide identity (17, 18). InvictusManeo is a K5 subcluster member, with a genome containing

Editor Catherine Putonti, Loyola University Chicago

Address correspondence to Christine A. Byrum, byrumc@cofc.edu.

The authors declare no conflict of interest.

Received 23 February 2022
Accepted 19 April 2022
Published 10 May 2022

TABLE 1 Characteristics of the InvictusManeo and Netyap bacteriophages

Parameter	Data for phage:	
	InvictusManeo	Netyap
GenBank accession no.	MZ958747	MW578835
SRA accession no.	SRX11158994	SRX11158997
Genome size (bp)	61,147	76,366
Collection location coordinates	35.316432N, 83.165618W	35.310051N, 83.187270W
GC content (%)	65.6	58.9
Coverage (×)	250	219
No. of predicted protein-coding genes	96	131
No. of tRNAs	1	12
No. of tmRNAs	0	0
Plaque size (mm) (*n* = 10)		
Range	1.9–4.7	1.3–2.5
Mean	3.0	1.9
Capsid size (nm) (*n* = 5 [InvictusManeo] or 6 [Netyap])		
Range	65–70	76–81
Mean	67.0	78.5
Tail length (nm) (*n* = 5 [InvictusManeo] or 6 [Netyap])		
Range	112–119	270–318
Mean	115.8	287.3

96 predicted protein-coding genes (36 with assigned putative functions), 3 orphan genes (*gp93* to *gp95*), and 1 tRNA gene. Genes include those for typical structural and assembly proteins, a lysis cassette (lysin A, lysin B, and holin), and the lysogeny-regulating proteins serine integrase (restricted to seven K5 subcluster members) and immunity repressor. Although their functional roles are unknown, gene products 1, 44, 50, 51, 54, and 92 may warrant further investigation; conserved in all K5 subcluster members, they are absent in other bacteriophages, which suggests that they may play crucial roles in the K5 subcluster. Whole-genome BLASTn alignment (11) of InvictusManeo to other bacteriophages indicates high levels of similarity to the K5 bacteriophages Collard (GenBank accession number NC_051593) (97.63% identity with 100% coverage), Kratio (GenBank accession number NC_028947) (99.02% identity with 97% coverage), and Larva (GenBank accession number NC_023724) (97.78% identity with 95% coverage).

The Netyap genome (L2 subcluster) contains 131 predicted protein-coding genes (55 with assigned putative functions), 2 orphan genes (*gp53* and *gp136*), and 12 tRNA genes.

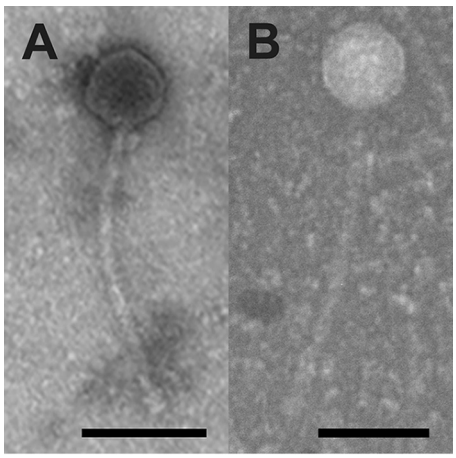

FIG 1 InvictusManeo (A) and Netyap (B) morphology examined using transmission electron microscopy. High-titer lysates placed on Formvar-coated copper grids were negatively stained with 1% uranyl acetate (2). Both phages exhibit *Siphoviridae* morphology. Scale bars, 100 nm.

Lysin A and lysin B are present, as well as the lysogeny-regulating proteins tyrosine integrase, immunity repressor, Cro, and excise, but holin was not detected. Gene products 21, 24, 25, 34, 48, and 51 (no known functions) may interest L2 phage investigators, because these occur only in the L2 subcluster and are conserved in all members. Although it is a temperate phage (lysogens can be readily isolated) (19), it is noteworthy that Netyap and most other L2 subcluster members form clear, lytic plaques. Finally, whole-genome BLASTn alignments (11) reveal a high level of nucleotide sequence conservation between Netyap and Faith1 (GenBank accession number NC_015584) (99.97% identity with 99% coverage).

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

ACKNOWLEDGMENTS

This study was kindly supported by the following: a SC INBRE program grant from the National Institutes of Health (grant P20GM103499-20), the HHMI-sponsored SEA-PHAGES program, and the College of Charleston Department of Biology. Also, sequencing and work performed by M.D.G. and J.R.W. were funded by a Western Carolina University Provost Grant.

Thanks go to Western Carolina University for sequencing and assembly of the genome and to Ken Grant of Wake Forest School of Medicine for imaging the viruses. Finally, we thank Alexander Cash and Sydney Winchel for finding and isolating the bacteriophages, Deborah Sera-Jacobs for reviewing the Netyap annotation, and Harrison Duckett, Stellamaris Onyechi, and Myah Renrick for assisting in the annotation of InvictusManeo.

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. Western Kentucky University Bioinformatics and Information Science Center, Bowling Green, KY. https://seaphages.org/media/docs/PECAAN_User_Guide_Dec7_2016.pdf.
7. Lukashin AV, 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.
8. 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.
9. Pacey M. 2016. Starterator guide. Pope W (ed). 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. 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.
12. 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.
13. 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.
14. 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.
15. 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.
16. 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.
17. Pope WH, Bowman CA, Russell DA, Jacobs-Sera D, Asai DJ, Cresawn SG, Jacobs WR, Hendrix RW, Lawrence JG, Hatfull GF. 2015. Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. Elife 4:e06416. https://doi.org/10.7554/eLife.06416.
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. Phagehunting Program. 2013. Phagehunting protocols: lysogeny experiments. https://phagesdb.org/media/workflow/protocols/pdfs/LysogenyProtocol_3.19.13.pdf.