AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Bacteriophages | Announcement

Complete genome sequences of cluster F1 and cluster B1 *Mycobacterium smegmatis* phages Karhdo and Basato

Jireh Kim,[1] Carlos Herrera,[1] Wai Yan Aung,[1] Gino Pablo Gonzales Boyles,[1] Carmina Chavez,[1] Mona Cibulka,[1] Emma Foley,[1] Jose Guerra,[1] Disha Byadarahalli Mohan Kumar,[1] Willow Levrant,[1] Lewis Lim,[1] Jose Llanes,[1] Zachary Kamuela O'Brien,[1] Art Pagaduan,[1] Justin Aaron Richardson,[1] Khristian Rosales,[1] James Schrecengost,[1] Tommy Shin,[1] Gabriel Strong-Lundquist,[1] Winnie Tat,[1] Fritz Vanderford,[1] Isabelle Vrinceanu,[1] Vicky Wang,[1] Stephanie Yang,[1] Christy Strong,[1] Philippos K. Tsourkas,[2] Kurt Regner[1]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT We present the complete genome sequences of *Mycobacterium smegmatis* phages Karhdo and Basato, isolated in Clark County, Nevada. The phages were isolated and annotated by students enrolled in undergraduate research courses over two semesters at the University of Nevada, Las Vegas.

KEYWORDS mycobacteria, bacteriophages, actinobacteriophage

Karhdo (36.131376 N, 115.240078 W) was isolated from compost, and Basato (35.985570 N, 115.123587 W) from a basil and tomato planter, both at private residences. Soil samples were incubated with enrichment broth, shaken (250 rpm, 2 h) at room temperature, followed by centrifugation and filter sterilization (0.22 µm) of the supernatant. Using *Mycobacterium smegmatis* mc^2 155 as the host, phages were considered pure after three rounds of plaque assays produced consistent plaque morphologies (1). Genomic DNA was isolated (Phage DNA Isolation Kit, Norgen Biotek), and samples were sequenced using the Illumina MiSeq System (v3 reagents) to yield 150 bp single-end reads with the reported coverage (Table 1). The total reads for Kardho and Basato were 1,205,997 and 210,277, respectively. The reads were quality trimmed and assembled *de novo* using Newbler (v. 2.9, 454 Life Science) to generate a single contig. Consed (v. 29) assessed accuracy, sequence completion, and phage genomic termini (2). For Transmission Electron Microscopy (TEM), 10 µLl of high titer (~10^{10}) lysate was added to copper 300 mesh grids (Ted Pella, Inc.) and stained with 1% phosphotungstate (Electron Microscopy Services). TEM was performed at 120 keV with a JEOL JEM-1400 Plus. Images were obtained with DigitalMicrograph with a Gata Orius SC1000 CCD camera.

The putative genes of Karhdo and Basato were identified from FASTA files using DNA Master (v5.38.8) and Phage Commander, which retrieves query results from Glimmer (v3.02b), GeneMark (v2.5), GeneMark.hmm (v3.25), GeneMarkS (v4.28), GeneMark with Heuristics (v3.25), GeneMarkS2, RAST (v2.0), MetaGene, and Aragorn for tRNAs (3–13). The annotation program Prokka (v1.14.6), which uses Prodigal (v2.6.3), was also used (14, 15). The putative genes and start codons were evaluated as described (16). The putative protein functions were assigned using Protein BLAST, CD-Search, and HHpred, with the E-value cutoffs of 10^{-7}, 0.001, and 0.001, respectively (17–19). Deep TMHMM (v1.0.24) and SOSUI were used to identify transmembrane domains (20, 21). Phage clusters and subclusters were determined with Phamerator (22), following protocols in reference (23). The default settings were used for all the software listed unless specified otherwise.

Kardho plaques were 0.5 cm in diameter, while Basato plaques varied in diameter from 0.04 to 0.4 cm. Karhdo and Basato both display siphovirus morphology with recorded capsid diameters and tail lengths of 83 nm and 316 nm for Karhdo, and 73 nm

Editor John J. Dennehy, Department of Biology, Queens College, Queens, New York, USA

Address correspondence to Kurt Regner, kurt.regner@unlv.edu.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 9 October 2023
Accepted 14 November 2023
Published 5 December 2023

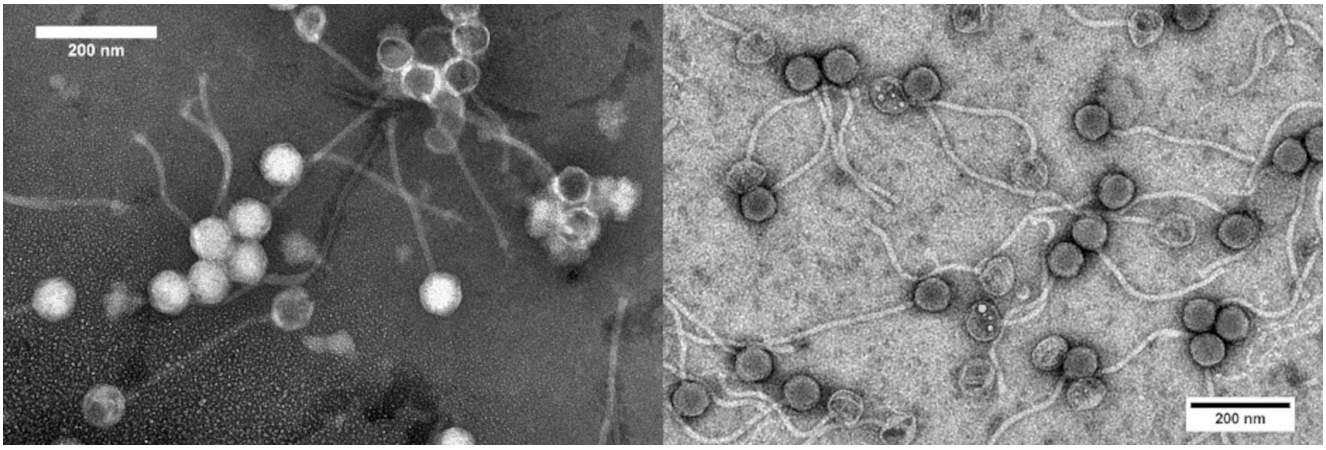

FIG 1 Transmission electron microscopy (TEM) of *Mycobacterium smegmatis* mc^2 155 bacteriophages Karhdo (left panel) and Basato (right panel). Samples were negatively stained with 10 mL of 1% phosphotungstate. Imaging was performed at 120 keV on JEOL-JEM-1400 Plus, Electron Microscopy Core Laboratory, University of Utah. Images captured by DigitalMicrograph with Gatan Orius SC1000 CCD camera.

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

Phage name	GenBank accession no.	SRA accession no.	Average coverage (×)	Cluster and subcluster	Genome length (bp)	GC content (%)	No. of genes
Basato	OR159661.1	SRX21748117	439×	B B1	68,518	66.5%	100
Karhdo	OR159669.1	SRX21748113	3421×	F F1	55,379	61.5%	100

and 306 nm for Basato, respectively (Fig. 1). Accession numbers and genome assembly results are listed in Table 1. Karhdo (cloudy temperate plaques), belongs to subcluster F1, while the Basato (clear lytic plaques) belongs to subcluster B1. Both phages have 100 genes. Karhdo has 41 genes with assigned protein functions, including a glycosyltransferase (genes 97 and 99), which has only been identified in cluster F phages. Basato has 29 genes with assigned protein functions. A programmed translational +1 frameshift in the tail assembly chaperone was found in Karhdo, and was annotated appropriately (3). Basato lacked a programmed translational frameshift. A complete list of genes and functions for both phages is available at The Actinobacteriophage Database (https://phagesdb.org).

ACKNOWLEDGMENTS

K.R. and C.S. acknowledge support from the HHMI SEA-PHAGES program and the Graham Hatfull Lab (University of Pittsburgh).

We thank Aude Picard (UNLV) for imaging the phages, and the TEM was supported by NSF EPSCoR RII Track-4 Award (OIA 2033286) and the UNLV School of Life Sciences. We thank Linda Nikolova and David Belnap for assistance at the Electron Microscopy Core Laboratory (University of Utah) and Julianne Grose (Brigham Young University) for the grid spotting protocol. In addition, we thank UNLV students Hector Aviles, Jessica Grifaldo, and Jocelyn Chavez Rios for their technical support.

K.R. acknowledges support from the National Institute of General Medical Sciences (GM103440; NV INBRE).

AUTHOR AFFILIATIONS

[1]School of Life Sciences, University of Nevada, Las Vegas, Nevada, USA
[2]School of Medicine and Public Health, University of Wisconsin, Madison, Wisconsin, USA

AUTHOR ORCIDs

Kurt Regner http://orcid.org/0000-0001-8540-5071

FUNDING

Funder	Grant(s)	Author(s)
HHS | NIH | National Institute of General Medical Sciences (NIGMS)	GM103440	Kurt Regner

AUTHOR CONTRIBUTIONS

Jireh Kim, Formal analysis, Writing – original draft | Carlos Herrera, Formal analysis, Writing – original draft | Wai Yan Aung, Formal analysis | Gino Pablo Gonzales Boyles, Formal analysis | Carmina Chavez, Formal analysis | Mona Cibulka, Formal analysis | Emma Foley, Formal analysis | Jose Guerra, Formal analysis | Disha Byadarahalli Mohan Kumar, Formal analysis | Willow Levrant, Formal analysis | Lewis Lim, Formal analysis | Jose Llanes, Formal analysis | Zachary Kamuela O'Brien, Formal analysis | Art Pagaduan, Formal analysis | Justin Aaron Richardson, Formal analysis, Investigation | Khristian Rosales, Formal analysis, Investigation | James Schrecengost, Formal analysis | Tommy Shin, Formal analysis | Gabriel Strong-Lundquist, Formal analysis | Winnie Tat, Formal analysis | Fritz Vanderford, Formal analysis | Isabelle Vrinceanu, Formal analysis | Vicky Wang, Formal analysis | Stephanie Yang, Formal analysis | Christy Strong, Project administration, Supervision, Writing – review and editing | Philippos K. Tsourkas, Methodology, Writing – review and editing | Kurt Regner, Conceptualization, Data curation, Formal analysis, Funding acquisition, Project administration, Supervision, Writing – review and editing

DATA AVAILABILITY

GenBank and SRA accession numbers are listed in Table 1.

REFERENCES

1. Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull G. 2018. Phage discovery guide. Howard hughes medical institute, Chevy Chase, MD. Available from: https://seaphagesdiscoveryguid.helpdocsonline.com/home
2. Gordon D, Abajian C, Green P. 1998. Consed: a graphical tool for sequence finishing. Genome Res 8:195–202. https://doi.org/10.1101/gr.8.3.195
3. Pope WH, Jacobs-Sera D. 2018. Annotation of bacteriophage genome sequences using DNA master: an overview, p 217–229. In Clokie MRJ, AM Kropinski, R Lavigne (ed), Bacteriophages. Springer, New York.
4. Lazeroff M, Ryder G, Harris SL, Tsourkas PK. 2021. Phage commander, an application for rapid gene identification in bacteriophage genomes using multiple programs. Phage2:204–213. https://doi.org/10.1089/phage.2020.0044
5. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. Bioinform 23:673–679. https://doi.org/10.1093/bioinformatics/btm009
6. Borodovsky M, McIninch J. 1993. GenMark: parallel gene recognition for both DNA strands. Comput Chem 17:123–133. https://doi.org/10.1016/0097-8485(93)85004-V
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. Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. implications for finding sequence motifs in regulatory regions. Nucleic Acids Res 29:2607–2618. https://doi.org/10.1093/nar/29.12.2607
9. Besemer J, Borodovsky M. 1999. Heuristic approach to deriving models for gene finding. Nucleic Acids Res 27:3911–3920. https://doi.org/10.1093/nar/27.19.3911
10. Lomsadze A, Gemayel K, Tang S, Borodovsky M. 2018. Modeling leaderless transcription and atypical genes results in more accurate gene prediction in prokaryotes. Genome Res 28:1079–1089. https://doi.org/10.1101/gr.230615.117
11. McNair K, Aziz RK, Pusch GD, Overbeek R, Dutilh BE, Edwards R. 2018. Phage genome annotation using the RAST pipeline, p 231–238. In Clokie MRJ, AM Kropinski, R Lavigne (ed), Bacteriophages. Springer, New York.
12. Noguchi H, Taniguchi T, Itoh T. 2008. Metageneannotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. DNA Res 15:387–396. https://doi.org/10.1093/dnares/dsn027
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. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinform 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153
15. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform 11:119. https://doi.org/10.1186/1471-2105-11-119
16. Salisbury A, Tsourkas PK. 2019. A method for improving the accuracy and efficiency of bacteriophage genome annotation. Int J Mol Sci 20:3391. https://doi.org/10.3390/ijms20143391
17. 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
18. Marchler-Bauer A, Bryant SH. 2004. CD-search: protein domain annotations on the fly. Nucleic Acids Res 32:W327–31. https://doi.org/10.1093/nar/gkh454
19. Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res 33:W244–8. https://doi.org/10.1093/nar/gki408
20. 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. Bioinform. Bioinform, Bioinformatics. https://doi.org/10.1101/2022.04.08.487609
21. Mitaku S, Hirokawa T, Tsuji T. 2002. Amphiphilicity index of polar amino acids as an aid in the characterization of amino acid preference at membrane-water interfaces. Bioinform 18:608–616. https://doi.org/10.1093/bioinformatics/18.4.608
22. Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF. 2011. Phamerator: a bioinformatic tool for comparative bacteriophage genomics. BMC Bioinform 12:395. https://doi.org/10.1186/1471-2105-12-395
23. Pope WH, Jacobs-Sera D, Russell DA, Peebles CL, Al-Atrache Z, Alcoser TA, Alexander LM, Bauerle CM, Bayles IM, Belfield KL, Best AA, Borjon A, Bowman CA, Boyer CA, Bradley KW, Bradley VA, Broadway LN, Budwal K, Busby KN, Campbell IW, Carey A, Caruso SM, Deng L, Guerra SL. 2011. Expanding the diversity of mycobacteriophages: insights into genome architecture and evolution. PLOS ONE 6:e16329. https://doi.org/10.1371/journal.pone.0016329