




Complete Genome Sequences of Mycobacteriophages Dallas and Jonghyun

Laricca Y. London,^a Mary A. Ayuk,^b Diana Effiom,^b Folasade Fashina,^b Briana J. Louis,^b Sara S. Tolsma,^c Adrian D. Allen,^b Leon A. Dickson,^b Somiranjana Ghosh,^b Ayele Gugssa,^b Hemayet Ullah,^b Glory B. Bassey,^b Lourds M. Fernando,^b Madison M. Moore,^b Jerome J. Oliver,^b Esohe G. Irabor,^b Swagata D. Roy,^b Benedict K. Quagrainie,^b Michael Smith,^b Howard University SEA-PHAGES Students,^b Winston A. Anderson,^b  Courtney J. Robinson^b

^aDepartment of Biological and Environmental Science, Alabama A&M University, Normal, Alabama, USA

^bDepartment of Biology, Howard University, Washington, DC, USA

^cDepartment of Biology, Northwestern College, Orange City, Iowa, USA

ABSTRACT Two temperate mycobacteriophages, Dallas and Jonghyun, were isolated from soil in Washington, DC, using the bacterial host *Mycobacterium smegmatis* mc²155. Analysis of the genomes revealed that Dallas and Jonghyun belong to clusters J and G, respectively. The structures of the genomes are typical of their respective clusters.

Mycobacteriophages Dallas and Jonghyun were isolated from two soil samples in Washington, DC (38.9227N, 77.0209W, and 38.9219N, 77.0186W, respectively), by students investigating mycobacteriophage diversity in the Phage Hunters Advancing Genomics and Evolutionary Science program at Howard University (1). Soil was incubated in 7H9 broth at 37°C for 2 h with shaking, followed by centrifugation and filtration (0.22- μ m-pore-size filters) of the supernatant. The filtrate was then incubated aerobically with *Mycobacterium smegmatis* mc²155 at 37°C for 48 h. This was followed by filtration of the enrichment and plating on *M. smegmatis* mc²155. Phage populations were subjected to three rounds of plaque purification, followed by DNA isolation using the Wizard DNA Clean-Up System (Promega) (2). Genomic DNA libraries were generated using NEBNext Ultra II FS kit v3 reagents with dual-indexed barcoding (New England BioLabs). Pooled libraries were run on an Illumina MiSeq sequencer to yield 150-bp single-end reads (Dallas, 1,058,535 bp; Jonghyun, 291,316 bp). The reads were trimmed and assembled using Newbler v2.9. Each case yielded a single contig, which was checked for completeness, accuracy, and phage genomic termini using Consed v29 (3–5). Each genome was annotated using the Phage Evidence Collection and Annotation Network (PECAAN; <http://pecaan.kbrinsgd.org>) and DNA Master v5.23.3 (cobamide2.bio.pitt.edu/computer.htm). Gene presence, start sites, and functions were determined using Glimmer v3.0, GeneMark v2.5, Starterator v1.1 (<https://github.com/SEA-PHAGES/starterator>), Phamerator Actino_Draft v402 (<https://phamerator.org>), BLAST v2.11.0+, PhagesDB, HHpred v3.0, and the NCBI Conserved Domain Database v3.19; tRNAs were detected using ARAGORN v1.2.38 and tRNAscan-SE v2.0 (6–13). Phylogeny was determined using the NCBI Sequence Read Archive (SRA) Taxonomy Analysis Tool (14). Default parameters were used for all software unless otherwise specified. Table 1 provides descriptions of each genome.

Both mycobacteriophages were identified as temperate based on plaque turbidity and the presence of genes associated with lysogeny, such as those that encode tyrosine integrase (SEA_DALLAS_80 and SEA_JONGHYUN_32), the excise protein (SEA_DALLAS_79 and SEA_JONGHYUN_34), and an immunity repressor (SEA_DALLAS_99 and SEA_JONGHYUN_33).

The double-stranded DNA genome of Dallas is linear with a 4-base overhang (ATCC). Actinobacteriophage clusters contain genomes with sequence similarity that is maintained for more than one-half of the length of each of the individual genomes and can be subdivided

Citation London LY, Ayuk MA, Effiom D, Fashina F, Louis BJ, Tolsma SS, Allen AD, Dickson LA, Ghosh S, Gugssa A, Ullah H, Bassey GB, Fernando LM, Moore MM, Oliver JJ, Irabor EG, Roy SD, Quagrainie BK, Smith M, Howard University SEA-PHAGES Students, Anderson WA, Robinson CJ. 2021. Complete genome sequences of mycobacteriophages Dallas and Jonghyun. Microbiol Resour Announc 10: e00304-21. <https://doi.org/10.1128/MRA.00304-21>.

Editor John J. Dennehy, Queens College CUNY

Copyright © 2021 London 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 Courtney J. Robinson, courtney.robinson@howard.edu.

Received 26 March 2021

Accepted 8 June 2021

Published 8 July 2021

TABLE 1 Genome characterization of two temperate mycobacteriophages

Phage name	GenBank accession no.	Cluster	Sequencing coverage (×)	Genome size (bp)	G+C content (%)	No. of protein-coding genes	No. of tRNAs
Dallas	MN062701	J	1,353	111,843	60.8	241	1
Jonghyun	MN444875	G	984	41,904	66.6	61	0

into subclusters (15, 16). Dallas was assigned to cluster J based on nucleotide similarity to other phages in this cluster (17). These phages are members of the *Siphoviridae* family (17). Similar to other cluster J phages, the genome length of Dallas is 111,843 bp, with one tRNA and 242 predicted protein-coding genes (17).

The genome of Jonghyun has a defined 11-base 3' overhang (CCCATGGCAT) and 98.1% of reads affiliated with *Siphoviridae*. The phage was assigned to cluster G, subcluster G1, based on nucleotide similarity to other subcluster G1 phages such as Chance64 (GenBank accession no. [MF919497](#)) (100.0%). Based on information in PhagesDB, the 41,904-bp genome size is typical of subcluster G1 (12).

Data availability. The complete genome sequences of phages Dallas and Jonghyun are available in GenBank (accession no. [MN062701](#) and [MN444875](#), respectively). The raw sequencing reads are available in the NCBI SRA under accession no. [SRX10013859](#) and [SRX10013862](#), respectively.

ACKNOWLEDGMENTS

This work was supported by the Howard Hughes Medical Institute (grant 52008098). This work was also supported by the Howard Hughes Medical Institute Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science program. Support was also provided by the Howard University Department of Biology.

We thank Rebecca A. Garlena, Deborah Jacobs-Sera, Welkin Pope, and Daniel A. Russell for genome sequencing, quality control, assembly, helpful discussions, and other assistance at the University of Pittsburgh. We also thank Howard University students Khukheper Awakowye, Naja Clarke, Carizma Forbes, Steven Harris, Sarah Solomon, Brooke Turner, and Syera Williams for initial annotation.

Howard University SEA-PHAGES student coauthors were Keishonna Belcher, Christina Charmant, Siara L. Davis, Siobhan Delaney, Somadinauwu Ezekwu, Cassidy Jones, Thomas Randle, Cymone Rice, and Trinita Sims.

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>.
- Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull GF. 2018. Phage discovery guide. Howard Hughes Medical Institute, Chevy Chase, MD.
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, 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 JB, 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>.
- 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.
- 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).
- Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23:673–679. <https://doi.org/10.1093/bioinformatics/btm009>.
- Besemer J, Borodovsky M. 2005. GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res* 33:W451–W454. <https://doi.org/10.1093/nar/gki487>.
- 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>.
- 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>.
- 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>.
- Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. *Bioinformatics* 33:784–786. <https://doi.org/10.1093/bioinformatics/btw711>.

13. 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>.
14. Katz KS, Shutov O, Lapoint R, Kimelman M, Brister JR, O'Sullivan C. 2021. A fast, scalable, MinHash-based k-mer tool to assess Sequence Read Archive next generation sequence submissions. *bioRxiv* 2021.02.16.431451. <https://doi.org/10.1101/2021.02.16.431451>.
15. Hatfull GF, Pedulla ML, Jacobs-Sera D, Cichon PM, Foley A, Ford ME, Gonda RM, Houtz JM, Hryckowian AJ, Kelchner VA, Namburi S, Pajcini KV, Popovich MG, Schleicher DT, Simanek BZ, Smith AL, Zdanowicz GM, Kumar V, Peebles CL, Jacobs WR, Lawrence JG, Hendrix RW. 2006. Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genet* 2:e92. <https://doi.org/10.1371/journal.pgen.0020092>.
16. 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, Wrynalek 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>.
17. Pope WH, Jacobs-Sera D, Best AA, Broussard GW, Connerly PL, Dedrick RM, Kremer TA, Offner S, Ogiefo AH, Pizzorno MC, Rockenbach K, Russell DA, Stowe EL, Stukey J, Thibault SA, Conway JF, Hendrix RW, Hatfull GF. 2013. Cluster J mycobacteriophages: intron splicing in capsid and tail genes. *PLoS One* 8:e69273. <https://doi.org/10.1371/journal.pone.0069273>.