Complete Genome Sequence of Cluster J Mycobacteriophage Superphikiman

Pratik Pradhan,[a] Sprikena Nako,[a] Trinh Tran,[a] Lavanya S. Aluri,[a] Dharman Anandarajan,[a] Niteesha Betini,[a] Shivangi D. Bhatt,[a] Swetha Chengalvala,[a] Nicole E. Cox,[a] Bela P. Delvadia,[a] Aishwary S. Desai,[a] Andrew M. Devaney,[a] Brenna K. Doyle,[a] Arden O. Edgerton,[a] Matthew C. Erlich,[a] Kevin C. Fitzpatrick,[a] Esha A. Gajjar,[a] Anjali Ganguly,[a] Ramnik S. Gill,[a] Pauline M. Good,[a] Nishtha Gupta,[a] Leila M. Haddad,[a] Esther J. Han,[a] Shelby Jain,[a] Andrew Jiang,[a] Andrew D. Jurgielewicz,[a] Devneet K. Kainth,[a] Jawhara M. Karam,[a] Mallika Kodavatiganti,[a] Sinja J. Kriete,[a] Catherine E. MacDonald,[a] Josh P. Maret,[a] Ashley E. Mathew,[a] Maanasa Natrajan,[a] Nusrat M. Nishu,[a] Nirali Patel,[a] Pooja D. Patel,[a] Shivani Patel,[a] Kaustav Patra,[a] Karima K. Rai,[a] Arghyadeep Sarkar,[a] Priyanka Shah,[a] Ravi K. Tata,[a] Andrew H. Tawfik,[a] Bhavya T. Thuremella,[a] Justina Toma,[a] Shika Veera,[a] Vamsee K. Vemulapalli,[a] Trevor V. Vidas,[a] Katy S. Vieira,[a] Gayathri Vijayakumar,[a] Tru A. Walor,[a] Clara R. White,[a] Brianna M. Wong,[a] Shu L. Zhao,[a] David W. Bollivar,[b] Matthew T. McDonald,[a] Ritu R. Dalia,[a] Kevin P. W. Smith,[a] Joy L. Little,[c] Susan M. R. Gurney[a]

[a]Department of Biology, Drexel University, Philadelphia, Pennsylvania, USA
[b]Department of Biological Sciences, Illinois Wesleyan University, Bloomington, Illinois, USA
[c]Department of Biological Sciences, North Carolina State University, Raleigh, North Carolina, USA

ABSTRACT Mycobacteriophage Superphikiman is a cluster J bacteriophage which was isolated from soil collected in Philadelphia, PA. Superphikiman has a 109,799-bp genome with 239 predicted genes, including 2 tRNA genes.

Mycobacteriophages are viruses which are known to infect mycobacterial host species (1). Mycobacteriophage Superphikiman was isolated from soil collected on the Drexel University campus in Philadelphia, PA, USA, by using *Mycobacterium smegmatis* mc²155 as a host. Superphikiman was isolated using the direct isolation procedure at 37°C in the fall of 2015. Its plaques were clear and consistently measured 1 mm in diameter. Electron microscopy showed that this bacteriophage has a *Siphoviridae* morphology.

Superphikiman plaques were selected, purified, and amplified, and the DNA was extracted. The Superphikiman genome was sequenced with the Illumina MiSeq platform using 150-bp single-end reads and had an average coverage of 335×. The genome was then annotated using the following databases and software: DNA Master (http://cobamide2.bio.pitt.edu), Glimmer (2), GeneMark (3), Starterator, Phamerator (4), PhagesDB.org, NCBI BLAST and Conserved Domain Database at NCBI (5, 6), HHPRED (7), ARAGORN (8), tRNAscan-SE (9), and PECAAN (http://pecaan.kbrinsgd.org).

Superphikiman is a member of actinobacteriophage cluster J (10). The Superphikiman genome is 109,799 bp in length, with a 4-base 3′ single-stranded DNA extension (right end 5′-ATCC) and a 61% G+C content. After annotation, 239 genes were identified, with 21.3% of genes being assigned a function. Superphikiman was compared to other bacteriophage genomes by using NCBI Blast (https://blast.ncbi.nlm.nih.gov/) and was found to be most similar to bacteriophage Courthouse (GenBank accession number JN698997), with 99% identity over 97% coverage.

The Superphikiman genome is arranged canonically, with the structural genes and viral assembly genes on the leftmost 35 kbp of the genome. At the rightmost 15 kbp of the genome, there are consecutive small reverse genes. A similar arrangement is found in other cluster J phages, including Courthouse. The Superphikiman genome has a lysis cassette consisting of lysin A, lysin B, and holin genes. Between the lysis cassette and immunity cassette there are 32 short reverse genes, which were not found to have

any specific functions. Given their position in the genome it is possible that these genes are involved in phage-host defense/counter defense systems. Other cluster J bacteriophages such as Courthouse have short reverse genes in this position; however, the number of genes present and the nucleotide sequences vary across this region (Phamerator.org). The immunity repressor in Superphikiman (gp191) is identical to the immunity repressor genes found in the cluster A1 bacteriophages Dynamix (gp75) (GenBank accession number AMD43071) and Abrogate (gp72) (GenBank accession number YP_009209461). The genome also contains the WhiB gene (gp126), which is shared by other cluster J bacteriophages but is not found in other phage clusters (Phamerator.org).

When the tRNA genes in Superphikiman were compared to those of three other cluster J phages (Courthouse, MiaZeal [GenBank accession number KM925136], and Ariel [GenBank accession number KM400683]), it was found that all four genomes contain 2 tRNA genes at approximately the same location in the genome and that these tRNA genes code for glycine and tyrosine. Other annotated cluster J bacteriophages (PhagesDB.org) have been found to contain tRNA genes for either glycine or tyrosine, or both, but not for other tRNA genes.

Accession number(s). This genome sequence has been deposited at GenBank under the accession number MF919534. The version described in this paper is the first version, MF919534.1.

REFERENCES

1. Hatfull GF. 2012. The secret lives of mycobacteriophages. Adv Virus Res 82:179–288. https://doi.org/10.1016/B978-0-12-394621-8.00015-7.
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.
3. 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.
4. 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.
5. 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.
6. 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.
7. Remmert M, Biegert A, Hauser A, Söding J. 2011. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods 9:173–175. https://doi.org/10.1038/nmeth.1818.
8. 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.
9. Lowe TM, Chan PP. 2016. tRNAscan-SE on-line: integrating search and contextual analysis of transfer RNA Genes. Nucleic Acids Res 44:W54–W57. https://doi.org/10.1093/nar/gkw413.
10. 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.