



Genome Sequences of Mycobacteriophages Jane and Sneeze, New Members of Cluster G

Catherine M. Mageeney,^a Cimrin Bhalla,^a Charles A. Bowman,^{b*} Bhavishya Devireddy,^a Adrienne P. Dzurick,^a Lee H. Graham,^a Marina Grossi,^a Margaret A. Kenna,^a Mikala A. Kowal,^a Drew D. Nielsen,^a Rachel A. Pallay,^a Kaitlyn M. Ruffing,^a Daniel A. Russell,^b Samantha L. Sarli,^{a*} Adama Shaw,^a Joseph W. Skibbens,^a Joseph N. Teyim,^{a*} Vassie C. Ware^a

Department of Biological Sciences, Lehigh University, Bethlehem, Pennsylvania, USA^a; Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, USA^b

ABSTRACT Jane and Sneeze are newly isolated phages of *Mycobacterium smegmatis* mc²155 from Hillsborough, NJ, and Palo Verde, Costa Rica, respectively. Both are cluster G, subcluster G1 mycobacteriophages. Notable nucleotide differences exist between genomes in the right half, including the presence of mycobacteriophage mobile element 1 (MPME1) in Jane.

A large number of mycobacteriophage genomes have been functionally annotated (1), allowing for insights into genomic structure, diversity, and genome evolution in this phage group. Further investigation of mycobacteriophage genomes and gene functions may aid in developing new molecular tools for diagnostics and therapeutics for tuberculosis (2, 3).

Mycobacteriophages Jane and Sneeze were isolated by enrichment from soil from Hillsborough, NJ, and Palo Verde, Costa Rica, respectively. Following phage purification and amplification, transmission electron microscopy revealed *Siphoviridae* morphologies. Double-stranded DNA (dsDNA) from each phage was isolated and sequenced by Ion Torrent, and reads were assembled into complete genomes using Newbler and Consed, with 3,162-fold and 2,606-fold coverage for Sneeze and Jane, respectively. The Sneeze genome is 42,429 bp, with 66.5% G+C content and 11-base 3' overhangs with the sequence 5'-TCCCATGGCAT. Jane has a 41,901-bp genome, 66.6% G+C content, and an 11-base 3' overhang of 5'-CCCCATGGCAT. The nucleotide difference at the 5' end of the Sneeze 3' overhang is not observed in other closely related mycobacteriophages.

Coding DNA sequence (CDS) predictions for both mycobacteriophages were determined using DNA Master (<http://cobamide2.bio.pitt.edu/>) embedded with GeneMark and Glimmer gene calling algorithms. Sneeze contains 64 predicted protein-coding genes, while Jane contains 63 predicted protein-coding genes. No tRNA genes or transfer-messenger (tmRNA) genes were predicted when analyzed with Aragorn and tRNAscan (4, 5). Functions were attributed to 31 of 64 Sneeze CDSs and 33 of 63 Jane CDSs, using BLASTP and HHPred (6). Sneeze and Jane are cluster G, subcluster G1 mycobacteriophages, bringing the total number of G1 subcluster members to 29 (1, 7, 8) (<http://www.phagesdb.org>). Jane is 99% identical to BPs (a G1 mycobacteriophage), spanning 99% of the genome. Sneeze is less identical (96%) to BPs, spanning 13% of the genome. Sneeze is most closely related to Taheera, with 97% identity spanning 31.7% of the genome. All but two genes (*gp32* and *gp33*) in Jane and Sneeze are rightward-transcribed.

Received 3 November 2016 Accepted 18 January 2017 Published 16 March 2017

Citation Mageeney CM, Bhalla C, Bowman CA, Devireddy B, Dzurick AP, Graham LH, Grossi M, Kenna MA, Kowal MA, Nielsen DD, Pallay RA, Ruffing KM, Russell DA, Sarli SL, Shaw A, Skibbens JW, Teyim JN, Ware VC. 2017. Genome sequences of mycobacteriophages Jane and Sneeze, new members of cluster G. *Genome Announc* 5:e01486-16. <https://doi.org/10.1128/genomeA.01486-16>.

Copyright © 2017 Mageeney 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 Vassie C. Ware, vcw0@lehigh.edu.

* Present address: Charles A. Bowman, The Scripps Research Institute, La Jolla, California, USA; Samantha L. Sarli, Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA; Joseph N. Teyim, OraSure Technologies, Inc., Bethlehem, Pennsylvania, USA.

The first halves of the Jane and Sneez genomes include CDSs predicted to encode proteins that function in structure and assembly, host integration, and lysis. These gene functions are consistent with those from other G1 subcluster phages and display syntenic relationships across all mycobacteriophage genomes. Many genes in the second half of the genome, as in other G1 phages, cannot currently be assigned a function. Putative functions in the right end of the genome include recombination factors and endonucleases. Additionally, Jane, but not Sneez, contains a mycobacteriophage mobile element 1 (MPME1) as *gp59*. MPME1 elements, found in many G cluster genomes, are described as “short-terminal inverted repeats” characteristic of other bacterial mobile elements (7). The MPME1 found in Jane is in phamily number 3572, which contains 100 members from various clusters. There are 12 additional G1 subcluster phages with the same MPME1 in the same location as in Jane, including BPs and Hope. Sneez is one of five G1 subcluster phages that lacks MPME1.

Accession number(s). The Jane and Sneez genome sequences are available in GenBank under accession numbers [KX588251](#) and [KX534004](#), respectively.

ACKNOWLEDGMENTS

This work was undertaken in the SEA-PHAGES courses at Lehigh University, with support provided by the Department of Biological Sciences and the Science Education Alliance of the Howard Hughes Medical Institute, Chevy Chase, MD.

We thank SEA-PHAGES team members, especially Welkin Pope, Deborah Jacobs-Sera, and Graham Hatfull, for program administration and thoughtful feedback.

REFERENCES

1. Pope WH, Bowman CA, Russell DA, Jacobs-Sera D, Asai DJ, Cresawn SG, Jacobs WR, Hendrix RW, Lawrence JG, Hatfull GF, Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science, Phage Hunters Integrating Research and Education, Mycobacterial Genetics Course. 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>.
2. 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>.
3. Hatfull GF. 2014. Mycobacteriophages: windows into tuberculosis. *PLoS Pathog* 10:e1003953. <https://doi.org/10.1371/journal.ppat.1003953>.
4. 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>.
5. 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/25.5.0955>.
6. 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>.
7. Sampson T, Broussard GW, Marinelli LJ, Jacobs-Sera D, Ray M, Ko CC, Russell D, Hendrix RW, Hatfull GF. 2009. Mycobacteriophages BPs, Angel and Halo: comparative genomics reveals a novel class of ultra-small mobile genetic elements. *Microbiology* 155:2962–2977. <https://doi.org/10.1099/mic.0.030486-0>.
8. Gan Y, Wu T, Liu P, Guo S. 2014. Characterization and classification of Bo4 as a cluster G mycobacteriophage that can infect and lyse *M. tuberculosis*. *Arch Microbiol* 196:209–218. <https://doi.org/10.1007/s00203-014-0954-6>.