



# Complete Genome Sequences of Cluster A Mycobacteriophages BobSwaget, Fred313, KADY, Lokk, MyraDee, Stagni, and StepMih

Kristen A. Butela,<sup>a</sup> Susan M. R. Gurney,<sup>b</sup>  Heather L. Hendrickson,<sup>c</sup> Janine M. LeBlanc-Straceski,<sup>d</sup> Anastasia M. Zimmerman,<sup>e</sup> Stephanie B. Conant,<sup>f</sup> Nikki E. Freed,<sup>c</sup> Olin K. Silander,<sup>c</sup> Joshua J. Thomson,<sup>g</sup> Charlotte A. Berkes,<sup>d</sup> Cristina Bertolez,<sup>d</sup> Courtney G. Davies,<sup>c</sup> Amber Elinsky,<sup>f</sup> Alison J. Hanlon,<sup>d</sup> Juliette Nersesyan,<sup>c</sup> Payal Patel,<sup>f</sup> John Sherwood,<sup>f</sup> Tiffany Tieu Ngo,<sup>f</sup> Kathryn A. Wisniewski,<sup>d</sup> Kathrine Yacoo,<sup>f</sup> Paul M. Arendse,<sup>c</sup> Nathan W. Bowlen,<sup>a</sup> Jasmina Cunmulaj,<sup>f</sup> Julie L. Downs,<sup>a</sup> Charlee A. Ferrenberg,<sup>a</sup> Alexandra E. Gassman,<sup>a</sup> Cody E. R. Gilligan,<sup>c</sup> Emily Gorkiewicz,<sup>f</sup> Christopher Harness,<sup>f</sup> Anthony Huffman,<sup>f</sup> Christina Jones,<sup>f</sup> Anna Julien,<sup>f</sup> Alexis E. Kupic,<sup>a</sup> Sayonara F. Latu,<sup>c</sup> Thomas J. Manning,<sup>c</sup> Danielle Maxwell,<sup>f</sup> Merrimack College SEA-PHAGES Annotators 2016, Catherine E. Meyer,<sup>c</sup> Madeleine Reardon,<sup>f</sup> Matthew Slaughter,<sup>a</sup> Royce Swasey,<sup>f</sup> Rebecca I. Tennent,<sup>c</sup> Victoria Torres,<sup>f</sup> Tamia Waller,<sup>f</sup> Rachel M. Worcester,<sup>a</sup> Brooke L. Yost,<sup>a</sup>  Steven G. Cresawn,<sup>i</sup> Rebecca A. Garlena,<sup>h</sup> Deborah Jacobs-Sera,<sup>h</sup>  Welkin H. Pope,<sup>h</sup> Daniel A. Russell,<sup>h</sup> Graham F. Hatfull,<sup>h</sup> Jacob D. Kagey<sup>f</sup>

Division of Natural and Health Sciences, Seton Hill University, Greensburg, Pennsylvania, USA<sup>a</sup>; Department of Biology, Drexel University, Philadelphia, Pennsylvania, USA<sup>b</sup>; Institute of Natural and Mathematical Sciences, Massey Phage Whānau, Massey University, Auckland, New Zealand<sup>c</sup>; Department of Biology, Merrimack College, North Andover, Massachusetts, USA<sup>d</sup>; Department of Biology, College of Charleston, Charleston, South Carolina, USA<sup>e</sup>; Biology Department, College of Engineering and Science, University of Detroit Mercy, Detroit, Michigan, USA<sup>f</sup>; Division of Integrated Biomedical Sciences, University of Detroit Mercy School of Dentistry, Detroit, Michigan, USA<sup>g</sup>; Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, USA<sup>h</sup>; Biology Department, James Madison University, Harrisonburg, Virginia, USA<sup>i</sup>

**ABSTRACT** Seven mycobacteriophages from distinct geographical locations were isolated, using *Mycobacterium smegmatis* mc<sup>2</sup>155 as the host, and then purified and sequenced. All of the genomes are related to cluster A mycobacteriophages, BobSwaget and Lokk in subcluster A2; Fred313, KADY, Stagni, and StepMih in subcluster A3; and MyraDee in subcluster A18, the first phage to be assigned to that subcluster.

**M**ycobacterium *smegmatis* mc<sup>2</sup>155 is a well-characterized actinobacterium that is used as a host for bacteriophage discovery (1). There are currently over 1,300 completely sequenced mycobacteriophages that have been discovered by participants in the Howard Hughes Medical Institute Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program (2, 3). These phages are considerably diverse, forming 24 clusters (clusters A to Z) and six singletons (those without close relatives). Cluster A is the largest cluster, with over 500 individual phage members, and it is subdivided into multiple subclusters based on overall sequence relationships (4, 5).

Seven mycobacteriophages were isolated from soil or compost samples using either enrichment culture or direct plating with the bacterial host *M. smegmatis* mc<sup>2</sup>155 at 26 to 37°C. All seven phages are morphologically members of the family *Siphoviridae*. The genomes were sequenced using the Illumina MiSeq platform with 150-bp reads and assembled using Newbler and Consed, with at least 300-fold coverage (6, 7). The

Received 21 September 2017 Accepted 27 September 2017 Published 26 October 2017

**Citation** Butela KA, Gurney SMR, Hendrickson HL, LeBlanc-Straceski JM, Zimmerman AM, Conant SB, Freed NE, Silander OK, Thomson JJ, Berkes CA, Bertolez C, Davies CG, Elinsky A, Hanlon AJ, Nersesyan J, Patel P, Sherwood J, Tieu Ngo T, Wisniewski KA, Yacoo K, Arendse PM, Bowlen NW, Cunmulaj J, Downs JL, Ferrenberg CA, Gassman AE, Gilligan CER, Gorkiewicz E, Harness C, Huffman A, Jones C, Julien A, Kupic AE, Latu SF, Manning TJ, Maxwell D, Merrimack College SEA-PHAGES Annotators 2016, Meyer CE, Reardon M, Slaughter M, Swasey R, Tennent RI, Torres V, Waller T, Worcester RM, Yost BL, Cresawn SG, Garlena RA, Jacobs-Sera D, Pope WH, Russell DA, Hatfull GF, Kagey JD. 2017. Complete genome sequences of cluster A mycobacteriophages BobSwaget, Fred313, KADY, Lokk, MyraDee, Stagni, and StepMih. *Genome Announc* 5:e01182-17. <https://doi.org/10.1128/genomeA.01182-17>.

**Copyright** © 2017 Butela 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 Jacob D. Kagey, [kageyja@udmercy.edu](mailto:kageyja@udmercy.edu).

**TABLE 1** Seven newly isolated cluster A mycobacteriophages

Phage name	Accession no.	Length (bp)	G+C content (%)	No. of ORFs <sup>a</sup>	No. of tRNAs	Cluster	Location of isolation
BobSwaget	<a href="#">MF185727</a>	50,400	63.3	89	1	A2	Andover, MA, USA
Lokk	<a href="#">MF324899</a>	51,008	63.4	88	1	A2	Andover, MA, USA
Fred313	<a href="#">MF373840</a>	50,053	64.0	85	2	A3	Chesterfield, MI, USA
KADY	<a href="#">MF185729</a>	50,898	64.2	90	3	A3	Warren, MI, USA
StepMih	<a href="#">MF185733</a>	50,841	64.0	93	3	A3	Auckland, New Zealand
Stagni	<a href="#">MF185732</a>	50,856	64.0	89	3	A3	Livonia, MI, USA
MyraDee	<a href="#">MF141539</a>	50,514	62.7	94	0	A18	Saltsburg, PA, USA

<sup>a</sup>ORFs, open reading frames.

genomes were annotated using DNA Master (<http://cobamide2.bio.pitt.edu>), Glimmer (8), GeneMark (9), Starterator, Phamerator (10), HHPRED (11), BLASTp searches against the NCBI nonredundant and actinobacteriophage (<http://phagesdb.org>) databases (12, 13), Aragorn (14), tRNAscanSE (15), and PECAAN (<http://pecaan.kbrinsgd.org>). Phage features are listed in Table 1.

All of the genomes are approximately 50 kb long and have nucleotide sequence similarities to those of cluster A phages. BobSwaget and Lokk are grouped into subcluster A2, and Fred313, KADY, Stagni, and StephMih are grouped into subcluster A3, according to their nucleotide sequence similarities. MyraDee is not closely related to phages in any particular subcluster and thus is the founding member of subcluster A18. We note that although Stagni and StephMih were isolated from geographically distinct locations, they share 99.8% nucleotide identity over their genome length. All of the phages have a typical cluster A genome organization, with the virion structure and assembly genes in their left arm and regulatory and replication functions in their right arm (16). However, they differ in the number and types of tRNA genes near their left genome end; Kady, Stagni, and StephMih have three tRNA genes (tRNA<sup>Asn</sup>, tRNA<sup>Leu</sup>, and tRNA<sup>Trp</sup>), Fred313 has two (tRNA<sup>Asn</sup>, tRNA<sup>Trp</sup>), BobSwaget and Lokk have one (tRNA<sup>Gln</sup>), and MyraDee has none.

All seven phages have features consistent with temperate lifestyles and encode putative immunity repressors related to L5 gp71 (17). However, they differ in the genes near the centers of the genomes that confer prophage maintenance. For example, Fred313, Stagni, and StephMih code for integrases of the tyrosine recombinase family, whereas KADY and MyraDee code for serine integrases (18). In contrast, BobSwaget and Lokk have *parABS* partitioning systems, as described for several other cluster A phages (19, 20), although their ParA and ParB proteins share only 62% and 48% amino acid sequence identities, respectively. Lokk, but not BobSwaget, codes for a putative RepA protein (gp36) that is implicated in extrachromosomal prophage replication and that shares 71% amino acid identity with phage CRB1 RepA (20). MyraDee codes for a putative Arda-like antirestriction protein (gp87).

**Accession number(s).** The genome sequences reported here have been deposited in GenBank under the accession numbers given in Table 1. The versions described here for these phages are the first versions reported.

## ACKNOWLEDGMENTS

Merrimack College SEA-PHAGES student contributors are listed at [https://seaphages.org/media/GA\\_authors/GA4\\_9-18-17.pdf](https://seaphages.org/media/GA_authors/GA4_9-18-17.pdf). We acknowledge Jung Min, Elizabeth Moak, and Alka Suresh for contributions to discovery and annotation.

This work was supported by the Howard Hughes Medical Institute (HHMI) SEA-PHAGES program and HHMI grant 54308198 to G.F.H. Support was also provided by NIH grants UL1GM118982, TL4GM118983, and RL5GM118981 to the University of Detroit Mercy ReBUILDetroit program.

## REFERENCES

- Hatfull GF. 2014. Molecular genetics of mycobacteriophages. *Microbiol Spectr* 2:1–36. <https://doi.org/10.1128/microbiolspec.MGM2-0032-2013>.
- Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. *Bioinformatics* 33:784–786. <https://doi.org/10.1093/bioinformatics/btw711>.
- Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, Den-

- nehy JJ, Denver DR, Dunbar D, Elgin SC, 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>.
4. 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>.
  5. Hatfull GF, Jacobs-Sera D, Lawrence JG, Pope WH, Russell DA, Ko CC, 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>.
  6. Miller JR, Koren S, Sutton G. 2010. Assembly algorithms for next-generation sequencing data. *Genomics* 95:315–327. <https://doi.org/10.1016/j.ygeno.2010.03.001>.
  7. Gordon D, Green P. 2013. Consed: a graphical editor for next-generation sequencing. *Bioinformatics* 29:2936–2937. <https://doi.org/10.1093/bioinformatics/btt515>.
  8. 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>.
  9. 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>.
  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. Söding J. 2005. Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21:951–960. <https://doi.org/10.1093/bioinformatics/bti125>.
  12. 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).
  13. Marchler-Bauer A, Panchenko AR, Shoemaker BA, Thiessen PA, Geer LY, Bryant SH. 2002. CDD: a database of conserved domain alignments with links to domain three-dimensional structure. *Nucleic Acids Res* 30: 281–283. <https://doi.org/10.1093/nar/30.1.281>.
  14. 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>.
  15. 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.
  16. 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>.
  17. Donnelly-Wu MK, Jacobs WR, Jr, Hatfull GF. 1993. Superinfection immunity of mycobacteriophage L5: applications for genetic transformation of mycobacteria. *Mol Microbiol* 7:407–417. <https://doi.org/10.1111/j.1365-2958.1993.tb01132.x>.
  18. Grindley ND, Whiteson KL, Rice PA. 2006. Mechanisms of site-specific recombination. *Annu Rev Biochem* 75:567–605. <https://doi.org/10.1146/annurev.biochem.73.011303.073908>.
  19. Dedrick RM, Mavrich TN, Ng WL, Cervantes Reyes JC, Olm MR, Rush RE, Jacobs-Sera D, Russell DA, Hatfull GF. 2016. Function, expression, specificity, diversity and incompatibility of actinobacteriophage parABS systems. *Mol Microbiol* 101:625–644. <https://doi.org/10.1111/mmi.13414>.
  20. Franceschelli JJ, Suarez CA, Terán L, Raya RR, Morbidoni HR. 2014. Complete genome sequences of nine mycobacteriophages. *Genome Announc* 2(3):e00181-14. <https://doi.org/10.1128/genomeA.00181-14>.