AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

VIRUSES

Genome Sequences of Three Cluster AU *Arthrobacter* Phages, Caterpillar, Nightmare, and Teacup

Tamarah L. Adair,[a] Emily Stowe,[b] Marie C. Pizzorno,[b] Gregory Krukonis,[b] Melinda Harrison,[c] Baylor University SEA-PHAGES,[a] Bucknell University SEA-PHAGES,[b] Cabrini University SEA-PHAGES,[c] Steven G. Cresawn,[d] Rebecca A. Garlena,[e] Daniel A. Russell,[e] Welkin H. Pope,[e] Deborah Jacobs-Sera,[e] Graham F. Hatfull[e]

Department of Biology, Baylor University, Waco, Texas, USA[a]; Department of Biology, Bucknell University, Lewisburg, Pennsylvania, USA[b]; Department of Biological Sciences, Cabrini University, Radnor, Pennsylvania, USA[c]; Department of Biology, James Madison University, Harrisonburg, Virginia, USA[d]; Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, USA[e]

ABSTRACT Caterpillar, Nightmare, and Teacup are cluster AU siphoviral phages isolated from enriched soil on *Arthrobacter* sp. strain ATCC 21022. These genomes are 58 kbp long with an average G+C content of 50%. Sequence analysis predicts 86 to 92 protein-coding genes, including a large number of small proteins with predicted transmembrane domains.

Received 11 September 2017 **Accepted** 11 October 2017 **Published** 9 November 2017

Citation Adair TL, Stowe E, Pizzorno MC, Krukonis G, Harrison M, Baylor University SEA-PHAGES, Bucknell University SEA-PHAGES, Cabrini University SEA-PHAGES, Cresawn SG, Garlena RA, Russell DA, Pope WH, Jacobs-Sera D, Hatfull GF. 2017. Genome sequences of three cluster AU *Arthrobacter* phages, Caterpillar, Nightmare, and Teacup. Genome Announc 5:e01121-17. https://doi.org/10.1128/genomeA.01121-17.

Address correspondence to Graham F. Hatfull, gfh@pitt.edu.

The *Actinobacteria* are a large and diverse group of soil bacteria with complex genetic relationships with each other and with their bacteriophages. An important genus of *Actinobacteria*, the *Arthrobacter*, includes common soil inhabitants that are important in biogeochemical cycling and bioremediation (1). Few *Arthrobacter* phages have been described relative to other *Actinobacteria* phages (2–4), and to better understand *Arthrobacter* phage diversity, students in the Science Education Alliance-Phage Hunters Advancing Genomic and Evolutionary Science (SEA-PHAGES) program used *Arthrobacter* sp. strain ATCC 20122 as a host to isolate and characterize bacteriophages from soil samples (5, 6). Here we report three newly discovered phages, Caterpillar, Nightmare, and Teacup, isolated on *Arthrobacter* sp. ATCC 20122, using enriched soil samples collected in Waco, TX, Chester, PA, and Lewisburg, PA, respectively. All three phages produce small clear plaques and have siphoviral virion morphologies with isometric heads 60 nm in diameter and flexible tails approximately 260 nm long.

Double-stranded DNA was extracted from high-titer phage lysates and sequenced on an Illumina MiSeq platform. Sequence reads from each genome were assembled into single contigs using Newbler and Consed (7), with minimum coverage of 160-fold. All genomes are members of cluster AU and have defined ends with 9-base complementary 3′ single-stranded DNA extensions (right end, 5′-CGCCGGCCT in Nightmare and Teacup and 5′-CGCCGGCCC in Caterpillar). The average G+C content for these three phages is 50.2%, which is 13.2% lower than the average G+C content of the bacterial host (8). All three phages are related to cluster AU phages, with greater than pairwise 82% identity spanning 65% of the genome lengths.

Genomes were annotated using DNAMaster (http://cobamide2.bio.pitt.edu), Glimmer (9), and GeneMark (10), and putative functions were assigned using BLASTP (11), HHPred (12), and Phamerator (13). All genes were transcribed in the forward direction, and no tRNAs were predicted by Aragorn (14). The number of predicted protein-coding genes ranges from 86 to 92, and up to 23% have putative functional assignments. All of the cluster AU phages have a lysis cassette near the left end of the genome. However, Caterpillar, Nightmare, and Teacup lack the putative glycosidase gene pres-

ent in CapnMurica and Gordon (15) and have an endolysin gene with predicted muramidase activity. No closely linked holin genes were identified. The virion structure and assembly genes include a fused capsid/capsid maturation protease, a conserved feature that is similarly found in cluster AM, BI, DJ, and CC phages, which infect *Arthrobacter*, *Streptomyces*, *Gordonia*, and *Rhodoccoccus*, respectively. Other genes shared among these diverse clusters code for terminase, portal, primase, ATP-dependent helicase, RecB-like exonuclease, and helix-turn-helix (HTH) DNA-binding domain proteins. No integrase or repressor genes were identified, and these phages are predicted to have lytic lifestyles. Interestingly, TMHMM predicted approximately 20 transmembrane proteins coded in each genome (13). These proteins are small (average size, 137 amino acids) and of unknown function. Most have a single transmembrane domain, although some (e.g., Caterpillar gp29) contain as many as five predicted transmembrane domains. Near their right ends, the genomes also have three HNH endonucleases with various sequence similarities.

Accession number(s). These phage genomes are available at GenBank with the accession no. MF140401, MF140423, and MF140432.

ACKNOWLEDGMENTS

This work was supported by Howard Hughes Medical Institute grant 54308198 to G.F.H.

We thank David Baltzegar, Erin Young, Kelly Sides, and Julie Macialek for sequencing of Nightmare at North Carolina State University. We thank the HHMI SEA-PHAGES program for support, and SEA-PHAGES students at Baylor University, Bucknell University, and Cabrini University (student contributors are listed at https://seaphages.org/media/GA_authors/GA3_Students_9-7-17.pdf).

REFERENCES

1. O'Loughlin EJ, Sims GK, Traina SJ. 1999. Biodegradation of 2-methyl, 2-ethyl, and 2-hydroxypyridine by an *Arthrobacter* sp. isolated from subsurface sediment. Biodegradation 10:93–104. https://doi.org/10.1023/A:1008309026751.
2. Klyczek KK, Bonilla JA, Jacobs-Sera D, Adair TL, Afram P, Allen KG, Archambault ML, Aziz RM, Bagnasco FG, Ball SL, Barrett NA, Benjamin RC, Blasi CJ, Borst K, Braun MA, Broomell H, Brown CB, Brynell ZS, Bue AB, Burke SO, Casazza W, Cautela JA, Chen K, Chimalakonda NS, Chudoff D, Connor JA, Cross TS, Curtis KN, Dahlke JA, Deaton BM, Degroote SJ, DeNigris DM, DeRuff KC, Dolan M, Dunbar D, Egan MS, Evans DR, Fahnestock AK, Farooq A, Finn G, Fratus CR, Gaffney BL, Garlena RA, Garrigan KE, Gibbon BC, Goedde MA, Guerrero Bustamante CA, Harrison M, Hartwell MC, Heckman EL. 2017. Tales of diversity: genomic and morphological characteristics of forty-six *Arthrobacter* phages. PLoS One 12:e0180517. https://doi.org/10.1371/journal.pone.0180517.
3. Šimoliūnas E, Kaliniene L, Stasilo M, Truncaitė L, Zajančkauskaitė A, Staniulis J, Nainys J, Kaupinis A, Valius M, Meškys R. 2014. Isolation and characterization of vB_ArS-ArV2—first *Arthrobacter* sp. infecting bacteriophage with completely sequenced genome. PLoS One 9:e111230. https://doi.org/10.1371/journal.pone.0111230.
4. Kaliniene L, Šimoliūnas E, Truncaitė L, Zajančkauskaitė A, Nainys J, Kaupinis A, Valius M, Meškys R. 2017. Molecular analysis of *Arthrobacter* myovirus vB_ArtM-ArV1: we blame it on the tail. J Virol 91. https://doi.org/10.1128/JVI.00023-17.
5. Hatfull GF. 2015. Innovations in undergraduate science education: going viral. J Virol 89:8111–8113. https://doi.org/10.1128/JVI.03003-14.
6. Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, Dennehy 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.
7. 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.
8. Russell DA, Hatfull GF. 2016. Complete genome sequence of *Arthrobacter* sp. ATCC 21022, a host for bacteriophage discovery. Genome Announc 4:e00168-16. https://doi.org/10.1128/genomeA.00168-16.
9. 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.
10. Borodovsky M, Lomsadze A. 2011. Gene identification in prokaryotic genomes, phages, metagenomes, and EST sequences with GeneMarkS suite. Curr Protoc Bioinform Chapter 4:Unit 4.5.1–4.5.17. https://doi.org/10.1002/0471250953.bi0405s35.
11. 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.
12. Söding J. 2005. Protein homology detection by HMM-HMM comparison. Bioinformatics 21:951–960. https://doi.org/10.1093/bioinformatics/bti125.
13. 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.
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. Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. Bioinformatics 33:784–786. https://doi.org/10.1093/bioinformatics/btw711.