AMERICAN SOCIETY FOR MICROBIOLOGY

Microbiology®
Resource Announcements

Genome Sequences of Five *Microbacterium foliorum* Phages, GaeCeo, NeumannU, Eightball, Chivey, and Hiddenleaf

Matthew D. Mastropaolo,[a] Patricia C. Fallest-Strobl,[a] DaeKwon M. Sequira,[a] Davia D. Campbell,[a] Christopher J. Negro,[a] Kho S. Tuang,[a] Ian M. Sigmund-Hamre,[a] Courtney L. Womack,[a] Thomas Mansbridge,[a] Amanda L. Metzler,[a] Emily G. Sasher,[a] Colleen Collins,[a] Nicholas K. Crowley,[a] Victoria R. Dower,[a] Megan Bates,[a] Christian Bjorkelo,[a] Hailey Johnson,[a] Lauren R. Salvitti,[b] Neumann University Phage Discovery Group

[a]Mathematics and Sciences, Neumann University, Aston, Pennsylvania, USA
[b]Charles River Laboratories, Newark, Delaware, USA

ABSTRACT Five siphoviruses were isolated from soil in southeastern Pennsylvania using *Microbacterium foliorum*. Bacteriophages NeumannU and Eightball have 25 predicted genes, Chivey and Hiddenleaf have 87 genes, and GaeCeo has 60 genes. Based on gene content similarity to sequenced actinobacteriophages, these five phages are distributed across clusters EA, EE, and EF.

Bacteriophages are an abundant and genetically diverse group of viruses that prey on bacteria (1). Here we report on the isolation and characterization of five bacteriophages isolated using *Microbacterium foliorum* NRRL B-24224.

All bacteriophages were isolated from soil collected in southeastern Pennsylvania (Table 1), using standard methods (2). Soil samples were suspended in peptone-yeast extract-calcium (PYCa) liquid medium and incubated with shaking at 250 rpm for 2 h at 30°C. The wash was then collected by centrifugation and filtered through a 0.22-μm filter, and the filtrate was plated in PYCa soft agar containing *M. foliorum* and incubated at 30°C for up to 48 h. Each bacteriophage was purified through three rounds of plating, and plaque morphologies are presented in Table 1. Bacteriophage morphology was determined by negative-staining transmission electron microscopy (TEM), and measurements of the capsids and tails were determined manually for a minimum of three particles. All isolated bacteriophages are siphoviruses (Table 1).

Genomic DNA was isolated from phage lysates using a $ZnCl_2$ precipitation method as previously described (2, 3). The DNA was prepared for sequencing using the NEBNext Ultra II FS kit (New England BioLabs) and sequenced using Illumina MiSeq (v3 reagents), yielding ~200,000 single-end 150-base reads. Untrimmed reads were assembled and then checked for completeness using Newbler v2.9 (4) and Consed v29 (5), respectively (6). Phages were assigned to clusters (Table 1) based on at least 35% gene content similarity to sequences in the actinobacteriophage database, phagesDB (7, 8).

Initial autoannotations of the genome were performed using DNA Master v5.23.6 (http://cobamide2.bio.pitt.edu/computer.htm) embedded with GeneMark v4.28 (9) and Glimmer v3.02b (10) and then refined using Phage Evidence Collection and Annotation Network v20211202 (PECAAN [https://pecaan.kbrinsgd.org/index.html]), Phamerator (11), and Starterator v462 (https://github.com/SEA-PHAGES/starterator). Transmembrane helices were predicted using TMHMM v2.0 (12), DeepTMHMM v1.0.11 (13), TOPCONS v2.0 (https://topcons.cbr.su.se/pred/) (14), and SOSUI v1.11 (15). tRNAs were predicted using ARAGORN v1.2.41 (16) and tRNAscanSE v2.0 (17), and functional assignments were made using BLASTP v2.9 (18) and HHPRED v3.2 (19). All annotations were performed with default parameters. Genome characteristics of

TABLE 1 Bacteriophage, plaque morphology, and genomic characteristics[a]

Phage name	Soil sample collection site	Isolation yr	Plaque morphology	Plaque size[b] (mm)	Capsid size[c] (nm)	Tail length[c] (nm)	Approx shotgun coverage (fold)	Genome length (bp)	Genome end characteristic	G+C content (%)	No. of ORFs[d]	No. of tRNAs	Cluster
Chivey	Garnet Valley, PA, 39.853972 N, 75.479417 W	2018	Clear middle, cloudy halo[e]	1–1.5	59–77	154–181	2,383	56,082	Circularly permutated	63.7	84	0	EF
Hiddenleaf	Garnet Valley, PA, 39.853972 N, 75.479417 W	2018	Small and clear	2	68	150–168	1,334	56,082	Circularly permutated	63.7	84	0	EF
Eightball	Chester, PA, 39.841972 N, 75.389932 W	2020	Small and clear	2	ND	ND	5,973	17,439	3′ single-stranded overhang, 5′-CCCGCCCCA-3′	68.7	25	0	EE
NeumannU	Aston, PA, 39.874167 N, 75.440889 W	2018	Small and clear	3–4	41	100	6,452	17,445	3′ single-stranded overhang, 5′-CCCGCCCCA-3′	68.7	25	0	EE
GaeCeo	Aston, PA, 39.871102 N, 75.436438 W	2018	Small and cloudy	0.9	50–58	134–146	2,726	40,168	Circularly permutated	63.4	60	1	EA[f]

[a] "ND" indicates the TEM was not performed.
[b] Based on the size of a minimum of 3 measured plaques.
[c] Based on the measurements of a minimum of 3 particles from a TEM.
[d] ORFs, open reading frames.
[e] Indicates a clear middle of the plaque with a diffuse or cloudy edge.
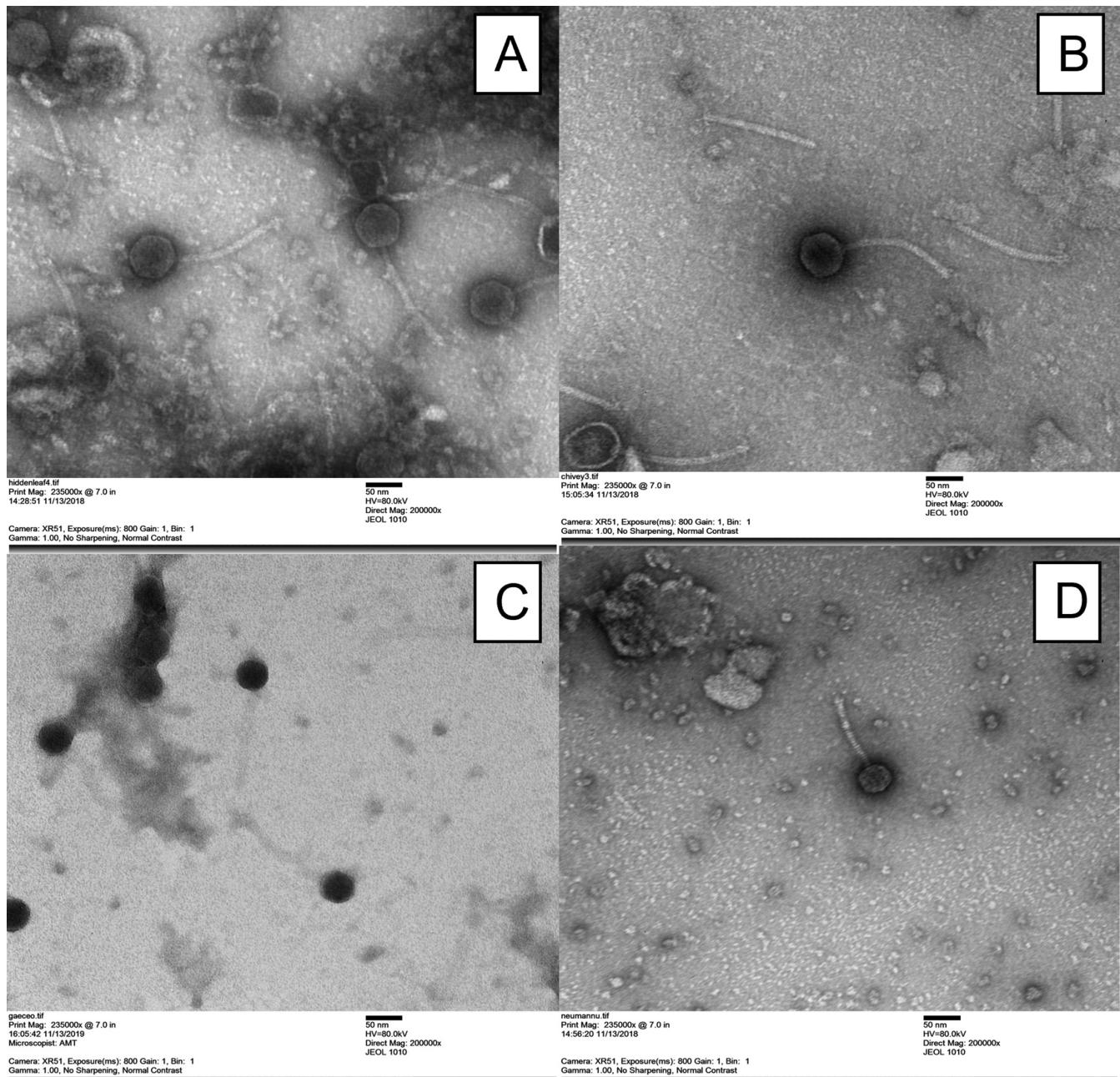[f] Subcluster EA9.

FIG 1 Transmission electron micrographs of bacteriophage morphology. (A) Hiddenleaf, cluster EF; (B) Chivey, cluster EF; (C) GaeCeo, cluster EA9; (D) NeumannU, cluster EE.

each bacteriophage are listed in Table 1, and the bacteriophage morphology is shown in Fig. 1.

NeumannU and Eightball are highly similar, sharing 99.9964% nucleotide identity, and contain 25 predicted genes, of which 22 are transcribed rightwards. The 3 genes that are transcribed leftward (genes 20 to 22) encode DNA-binding proteins. Hiddenleaf and Chivey also share 99.5% nucleotide identity, with all 84 predicted genes transcribed rightwards. GaeCeo has 60 predicted genes, including 1 tRNA tRNAPro, with structure, assembly, and lysis genes occupying the left half of the genome (genes 2 to 27) and transcribed rightwards and DNA metabolism genes (genes 34 to 51) occupying the right half of the genome and transcribed leftwards, with the exception of the rightmost gene (gene 60), which is transcribed rightwards. None of the five phages encode identifiable immunity repressor or

integrase functions, and they are therefore likely to be lytic, consistent with the life cycle of other phages in these clusters.

Data availability. All genomes, NeumannU, Eightball, Chivey, Hiddenleaf, and GaeCeo, are available at GenBank under accession no. MT657332, OK040783, MT684591, MN497954, and MT657343 and Sequence Read Archive (SRA) no. SRX15940725, SRX15940721, SRX15940720, SRX15940723, and SRX15940722, respectively.

REFERENCES

1. Pedulla ML, Ford ME, Houtz JM, Karthikeyan T, Wadsworth C, Lewis JA, Jacobs-Sera D, Falbo J, Gross J, Pannunzio NR, Brucker W, Kumar V, Kandasamy J, Keenan L, Bardarov S, Kriakov J, Lawrence JG, Jacobs WR, Hendrix RW, Jr, Hatfull GF. 2003. Origins of highly mosaic mycobacteriophage genomes. Cell 113:171–182. https://doi.org/10.1016/S0092-8674(03)00233-2.
2. Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull G. 2018. Phage discovery guide. Howard Hughes Medical Institute, Chevy Chase, MD. https://seaphagesphagediscoveryguide.helpdocsonline.com/home.
3. Santos MA. 1991. An improved method for the small scale preparation of bacteriophage DNA based on phage precipitation by zinc chloride. Nucleic Acids Res 19:5442. https://doi.org/10.1093/nar/19.19.5442.
4. 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.
5. Gordon D, Green P. 2013. Consed: a graphical editor for next-generation sequencing. Bioinformatics 29:2936–2937. https://doi.org/10.1093/bioinformatics/btt515.
6. 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.
7. Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. Bioinformatics 33:784–786. https://doi.org/10.1093/bioinformatics/btw711.
8. Pope WH, Mavrich TN, Garlena RA, Guerrero-Bustamante CA, Jacobs-Sera D, Montgomery MT, Russell DA, Warner MH, Hatfull GF, Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES). 2017. Bacteriophages of Gordonia spp. display a spectrum of diversity and genetic relationships. mBio 8:e01069-17. https://doi.org/10.1128/mBio.01069-17.
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. 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.
11. 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.
12. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567–580. https://doi.org/10.1006/jmbi.2000.4315.
13. Hallgren J, Tsirigos KD, Pedersen MD, Almagro AJ, Marcatili P, Nielsen H, Krogh A, Winther O. 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. bioRxiv. https://doi.org/10.1101/2022.04.08.487609.
14. Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic Acids Res 43:W401–W407. https://doi.org/10.1093/nar/gkv485.
15. Hirokawa T, Boon-Chieng S, Mitaku S. 1998. SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics 14:378–379. https://doi.org/10.1093/bioinformatics/14.4.378.
16. 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.
17. Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res 44:W54–W57. https://doi.org/10.1093/nar/gkw413.
18. 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.
19. 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.