AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Klebsiella pneumoniae* Siphophage Sugarland

Samuel G. Erickson,[a] Lauren Lessor,[a] Chandler J. O'Leary,[a] Jason J. Gill,[a] Mei Liu[a]

[a]Center for Phage Technology, Texas A&M University, College Station, Texas, USA

ABSTRACT *Klebsiella pneumoniae* is a Gram-negative bacterium associated with the gastrointestinal tract and is a significant nosocomial pathogen due to its antibiotic resistance. Phage therapy against *K. pneumoniae* may prove useful in treating infections caused by this bacterium. This announcement describes the genome of the T5-like *K. pneumoniae* siphophage Sugarland.

Klebsiella pneumoniae is a Gram-negative bacterium found in soil and the mucosal lining of the intestinal tract. It can cause pneumonia, urinary tract infections, sepsis, and soft tissue infections and is a significant nosocomial pathogen due to its resistance to antibiotics (1, 2). Carbapenemase-producing strains of sequence type 258 (ST258) are among the most prevalent in U.S. clinical centers (3). *K. pneumoniae* phage Sugarland was isolated from a wastewater treatment plant in College Station, Texas, in October 2016 using a carbapenem-resistant *K. pneumoniae* ST258 clinical isolate as the host. Upon isolation, it was identified as a siphophage using negative-stain transmission electron microscopy performed at the Texas A&M University Microscopy and Imaging Center.

Phage genomic DNA was prepared as described previously and sequenced on the Illumina MiSeq platform as paired-end 250-bp reads (4). FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to quality control reads, and reads were trimmed with the FastX Toolkit (hannonlab.cshl.edu) before being assembled to a single contig at 103.3-fold coverage using SPAdes 3.5.0 (5). Contig completion was confirmed by PCR and sequencing of the resulting product. Along with manual correction, Glimmer3 (6) and MetaGeneAnnotator (7) were used to predict protein-coding genes; tRNA genes were predicted with ARAGORN (8). Sequence similarity searches by BLASTp (9) and conserved domain searches with InterProScan 5 (10) were used to predict protein functions. All analyses were conducted via the CPT Galaxy (11) and WebApollo (12) interfaces (cpt.tamu.edu) using default parameters.

The 111,103-bp double-stranded DNA genome of phage Sugarland has a coding density of 87% and a GC content of 45%, which is significantly lower than the 58% GC content of the host (13). Analysis showed 174 predicted protein-coding genes and 24 identified tRNA genes. The progressiveMauve algorithm (14) was used to compare Sugarland's nucleotide similarity against the NR database, and the most similar organism at 78% sequence identity was the *Klebsiella* phage vB_Kpn_IME260 (GenBank accession no. KX845404). BLASTp analysis of the Sugarland proteome showed close homology to other T5-like phages, including the canonical phage T5 itself, with 110 similar proteins (E value < 0.001). Analysis by PhageTerm (15) was unable to precisely determine the extent of the terminal repeats typically associated with T5-like phages, and this genome was reopened to be syntenic to T5 with the predicted *dmp* as the first gene of the genome.

The genome displayed a 1,587-bp noncoding region characteristic of T5-like phages (16). The structural tail fiber and side tail fiber genes were identified, including the tail tip, baseplate, major tail subunit, and L-shaped side tail fiber proteins. Similar to T5, the tape measure chaperone protein of Sugarland did not contain a predicted frameshift

Received 26 July 2018 **Accepted** 2 October 2018 **Published** 15 November 2018

Citation Erickson SG, Lessor L, O'Leary CJ, Gill JJ, Liu M. 2018. Complete genome sequence of *Klebsiella pneumoniae* siphophage Sugarland. Microbiol Resour Announc 7:e01014-18. https://doi.org/10.1128/MRA.01014-18.

Editor Jason Stajich, University of California, Riverside

Address correspondence to Mei Liu, meiliu@tamu.edu.

signal (17). Genes involved in lysis, including a holin, a D-alanyl-D-alanine carboxypeptidase endolysin, and a partially embedded two-component spanin complex, were identified (18). There were 3 HNH endonucleases identified in the Sugarland genome, but all had free-standing open reading frames (ORFs) and were not introns (19). Interestingly, an NAD^+-dependent protein deacetylase in the sirtuin-2 family was found. Present in many organisms, protein acetylation helps regulate protein-protein interaction, DNA-protein interaction, and protein stability (20).

Data availability. The genome sequence of phage Sugarland was deposited under GenBank accession no. MG459987 and BioSample accession no. SAMN10128164. The BioProject accession number is PRJNA222858, and the SRA accession number is SRR7902581.

ACKNOWLEDGMENTS

This work was primarily supported by funding from award numbers EF-0949351 and DBI-1565146 from the National Science Foundation. Additional support came from the Center for Phage Technology (CPT), an Initial University Multidisciplinary Research Initiative supported by Texas A&M University and Texas AgriLife, and from the Department of Biochemistry and Biophysics.

We thank Thomas Walsh, Weill Cornell Medical School, for the provision of bacterial isolates. We are grateful for the advice and support of the CPT staff. This announcement was prepared in partial fulfillment of the requirements for BICH464 Phage Genomics, an undergraduate course at Texas A&M University.

REFERENCES

1. Thaden JT, Lewis SS, Hazen KC, Huslage K, Fowler VG, Jr, Moehring RW, Chen LF, Jones CD, Moore ZS, Sexton DJ, Anderson DJ. 2014. Rising rates of carbapenem-resistant *Enterobacteriaceae* in community hospitals: a mixed-methods review of epidemiology and microbiology practices in a network of community hospitals in the southeastern United States. Infect Control Hosp Epidemiol 35:978–983. https://doi.org/10.1086/677157.
2. Podschun R, Ullmann U. 1998. *Klebsiella* spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. Clin Microbiol Rev 11:589–603.
3. Satlin MJ, Chen L, Patel G, Gomez-Simmonds A, Weston G, Kim AC, Seo SK, Rosenthal ME, Sperber SJ, Jenkins SG, Hamula CL, Uhlemann AC, Levi MH, Fries BC, Tang YW, Juretschko S, Rojtman AD, Hong T, Mathema B, Jacobs MR, Walsh TJ, Bonomo RA, Kreiswirth BN. 2017. Multicenter clinical and molecular epidemiological analysis of bacteremia due to carbapenem-resistant enterobacteriaceae (CRE) in the CRE epicenter of the United States. Antimicrob Agents Chemother 61:e02349-16. https://doi.org/10.1128/AAC.02349-16.
4. Gill JJ, Berry JD, Russell WK, Lessor L, Escobar-Garcia DA, Hernandez D, Kane A, Keene J, Maddox M, Martin R, Mohan S, Thorn AM, Russell DH, Young R. 2012. The *Caulobacter crescentus* phage phiCbK: genomics of a canonical phage. BMC Genomics 13:542. https://doi.org/10.1186/1471-2164-13-542.
5. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
6. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res 27:4636–4641. https://doi.org/10.1093/nar/27.23.4636.
7. Noguchi H, Taniguchi T, Itoh T. 2008. MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. DNA Res 15:387–396. https://doi.org/10.1093/dnares/dsn027.
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. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421.
10. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30:1236–1240. https://doi.org/10.1093/bioinformatics/btu031.
11. Cock PJ, Gruning BA, Paszkiewicz K, Pritchard L. 2013. Galaxy tools and workflows for sequence analysis with applications in molecular plant pathology. PeerJ 1:e167. https://doi.org/10.7717/peerj.167.
12. Lee E, Helt GA, Reese JT, Munoz-Torres MC, Childers CP, Buels RM, Stein L, Holmes IH, Elsik CG, Lewis SE. 2013. Web Apollo: a Web-based genomic annotation editing platform. Genome Biol 14:R93. https://doi.org/10.1186/gb-2013-14-8-r93.
13. Hua X, Chen Q, Li X, Feng Y, Ruan Z, Yu Y. 2014. Complete genome sequence of *Klebsiella pneumoniae* sequence type 17, a multidrug-resistant strain isolated during tigecycline treatment. Genome Announc 2:e01337-14. https://doi.org/10.1128/genomeA.01337-14.
14. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5:e11147. https://doi.org/10.1371/journal.pone.0011147.
15. Garneau JR, Depardieu F, Fortier LC, Bikard D, Monot M. 2017. PhageTerm: a tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. Sci Rep 7:8292. https://doi.org/10.1038/s41598-017-07910-5.
16. Davison J. 2015. Pre-early functions of bacteriophage T5 and its relatives. Bacteriophage 5:e1086500. https://doi.org/10.1080/21597081.2015.1086500.
17. Zivanovic Y, Confalonieri F, Ponchon L, Lurz R, Chami M, Flayhan A, Renouard M, Huet A, Decottignies P, Davidson AR, Breyton C, Boulanger P. 2014. Insights into bacteriophage T5 structure from analysis of its morphogenesis genes and protein components. J Virol 88:1162–1174. https://doi.org/10.1128/JVI.02262-13.
18. Summer EJ, Berry J, Tran TA, Niu L, Struck DK, Young R. 2007. *Rz*/*Rz1* lysis gene equivalents in phages of Gram-negative hosts. J Mol Biol 373:1098–1112. https://doi.org/10.1016/j.jmb.2007.08.045.
19. Belfort M, Bonocora RP. 2014. Homing endonucleases: from genetic anomalies to programmable genomic clippers. Methods Mol Biol 1123:1–26. https://doi.org/10.1007/978-1-62703-968-0_1.
20. North BJ, Verdin E. 2004. Sirtuins: Sir2-related NAD-dependent protein deacetylases. Genome Biol 5:224. https://doi.org/10.1186/gb-2004-5-5-224.