AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

# Complete Genome Sequence of *Klebsiella pneumoniae* Siphophage Sanco

**Ryan W. Richardson,[a] Lauren Lessor,[a] Chandler O'Leary,[a] Jason Gill,[a] Mei Liu[a]**

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

**ABSTRACT** *Klebsiella pneumoniae* is an opportunistic pathogen that is the cause of several hospital-acquired infections. Bacteriophages that target this bacterium could be used therapeutically as novel antimicrobial agents. Here, we present the complete genome sequence of the T1-like *K. pneumoniae* phage Sanco.

*Klebsiella pneumoniae*, especially with the continued emergence of multidrug-resistant strains, has become a significant public health threat in hospital settings (1–3). Characterization of *K. pneumoniae* phages may prove useful in developing new treatments for infections caused by this bacterium.

Phage Sanco was isolated in 2013 from a wastewater treatment plant in College Station, TX, against a deidentified *K. pneumoniae* clinical isolate. *K. pneumoniae* was cultured on tryptic soy broth or agar (Difco) at 37°C with aeration. Phage were cultured and propagated by the soft-agar overlay method (4). It was identified as a siphophage using negative-staining transmission electron microscopy, performed at the Texas A&M University Microscopy and Imaging Center as described previously (5). Phage genomic DNA was prepared using a modified Promega Wizard DNA cleanup kit protocol (5). Pooled indexed DNA libraries were prepared using the Illumina TruSeq Nano LT kit, and the sequence was obtained with the Illumina MiSeq platform using the MiSeq v2 500-cycle reagent kit following the manufacturer's instructions, producing 1,296,046 paired-end reads for the index containing the phage Sanco genome. FastQC v0.11.5 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to quality control reads. The reads were trimmed with FastX Toolkit v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/download.html) before being assembled using SPAdes v3.5.0 (6). Contig completion was confirmed by PCR using primers 5′-CCGGTTTGTCGATATCATCC-3′ and 5′-ACGGAGGTGTTTTCAATCCA-3′ facing off the ends of the assembled contig and Sanger sequencing of the resulting product, with the contig sequence manually corrected to match the resulting Sanger sequencing read. GLIMMER v3.0 (7) and MetaGeneAnnotator 1.0 (8) were used to predict protein coding genes with manual verification, and tRNA genes were predicted with ARAGORN 2.36 (9). Rho-independent termination sites were identified via TransTermHP (http://transterm.cbcb.umd.edu/). Sequence similarity searches were done by using BLASTp v2.2.28 (10) against the NCBI nonredundant (nr) and UniProt Swiss-Prot (11) and TrEMBL databases. InterProScan 5.15-54.0 (12), LipoP (13), and TMHMM v2.0 (14) were used to predict protein function. HHpred with ummiclust30_2018_08 for multiple sequence alignment (MSA) generation and PDB_mmCIF70 for modeling in the HHsuite v3.0 release were also used for functional prediction (15). All analyses were conducted using default settings via the Center for Phage Technology (CPT) Galaxy (16) and WebApollo (17) interfaces (https://cpt.tamu.edu/galaxy-pub).

Sanco was assembled at 28-fold coverage into a complete contig of 48,790 bp. The GC content of the genome is 51%, which is lower than that of the host (57%) (18). Determined by BLASTn against the NCBI nucleotide (nt) database, Sanco shares greater

**Citation** Richardson RW, Lessor L, O'Leary C, Gill J, Liu M. 2019. Complete genome sequence of *Klebsiella pneumoniae* siphophage Sanco. Microbiol Resour Announc 8:e01252-19. https://doi.org/10.1128/MRA.01252-19.

**Editor** Simon Roux, DOE Joint Genome Institute



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

**Received** 3 October 2019
**Accepted** 13 October 2019
**Published** 31 October 2019

than 83% overall nucleotide identity (E value = 0) with a group of characterized T1-like *Klebsiella* phages, including Sushi (GenBank accession no. KT001920) (19), Skenny (MK931444) (20), Sweeny (MK931443) (21), and Shelby (MK931445) (22). The Sanco genome was opened to be syntenic with those of phage T1 (NC_005833) and phage TLS (NC_009540). Sanco proteins sharing homology (determined by BLASTp search against the NCBI nr database at an E value cutoff of $10^{-3}$) with T1 proteins include those involved in phage morphogenesis and DNA replication. A full lysis cassette was identified in the Sanco genome, and it included a holin, signal-arrest-release (SAR) endolysin, and unimolecular spanin. The endolysin is predicted to have a glycosidase activity.

**Data availability.** The genome sequence of phage Sanco was submitted to GenBank under accession no. MK618657. The associated BioProject, SRA, and BioSample accession numbers are PRJNA222858, SRR8772108, SAMN11236500, respectively.

## ACKNOWLEDGMENTS

This work was supported by funding from the National Science Foundation (award no. EF-0949351 and DBI-1565146), and by the National Institutes of Health (NIAID award no. AI121689). 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 at Texas A&M University.

We thank Karen Frank at the NIH Clinical Center for providing bacterial isolates. We are grateful for the advice and support of the CPT staff and the Texas A&M University Microscopy and Imaging Center.

This announcement was prepared in partial fulfillment of the requirements for BICH464 Phage Genomics, an undergraduate course at Texas A&M University.

## REFERENCES

1. Struve C, Krogfelt KA. 2004. Pathogenic potential of environmental *Klebsiella pneumoniae* isolates. Environ Microbiol 6:584–590. https://doi.org/10.1111/j.1462-2920.2004.00590.x.
2. Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, Jenney A, Connor TR, Hsu LY, Severin J, Brisse S, Cao H, Wilksch J, Gorrie C, Schultz MB, Edwards DJ, Nguyen KV, Nguyen TV, Dao TT, Mensink M, Minh VL, Nhu NT, Schultsz C, Kuntaman K, Newton PN, Moore CE, Strugnell RA, Thomson NR. 2015. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. Proc Natl Acad Sci U S A 112:E3574–E3581. https://doi.org/10.1073/pnas.1501049112.
3. Arnold RS, Thom KA, Sharma S, Phillips M, Kristie Johnson J, Morgan DJ. 2011. Emergence of *Klebsiella pneumoniae* carbapenemase-producing bacteria. South Med J 104:40–45. https://doi.org/10.1097/SMJ.0b013e3181fd7d5a.
4. Adams MK. 1959. Bacteriophages. Interscience Publishers, Inc., New York, NY.
5. 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.
6. 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.
7. 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.
8. 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.
9. 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.
10. 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.
11. The UniProt Consortium. 2018. UniProt: the universal protein knowledgebase. Nucleic Acids Res 46:2699. https://doi.org/10.1093/nar/gky092.
12. 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.
13. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci 12:1652–1662. https://doi.org/10.1110/ps.0303703.
14. 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.
15. Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, Lozajic M, Gabler F, Soding J, Lupas AN, Alva V. 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J Mol Biol 430:2237–2243. https://doi.org/10.1016/j.jmb.2017.12.007.
16. 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.
17. 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.
18. Liu P, Li P, Jiang X, Bi D, Xie Y, Tai C, Deng Z, Rajakumar K, Ou HY. 2012. Complete genome sequence of *Klebsiella pneumoniae* subsp.

*pneumoniae* HS11286, a multidrug-resistant strain isolated from human sputum. J Bacteriol 194:1841–1842. https://doi.org/10.1128/JB.00043-12.
19. Nguyen DT, Lessor LE, Cahill JL, Rasche ES, Kuty Everett GF. 2015. Complete genome sequence of *Klebsiella pneumoniae* carbapenemase-producing *K. pneumoniae* siphophage Sushi. Genome Announc 3:e00994-15. https://doi.org/10.1128/genomeA.00994-15.
20. Gramer J, Kenny S, Newkirk H, Liu M, Gill JJ, Ramsey J. 2019. Complete genome sequence of *Klebsiella pneumoniae* siphophage Skenny. Microbiol Resour Announc 8:e01036-19. https://doi.org/10.1128/MRA.01036-19.
21. Martinez N, Williams E, Newkirk H, Liu M, Gill JJ, Ramsey J. 2019. Complete genome sequence of *Klebsiella pneumoniae* phage Sweeny. Microbiol Resour Announc 8:e01047-19. https://doi.org/10.1128/MRA.01047-19.
22. Saldana R, Newkirk H, Liu M, Gill JJ, Ramsey J. 2019. Complete genome sequence of Shelby, a siphophage infecting carbapenemase-producing *Klebsiella pneumoniae*. Microbiol Resour Announc 8:e01037-19. https://doi.org/10.1128/MRA.01037-19.