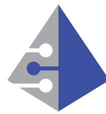

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Klebsiella pneumoniae* Myophage Mineola

Justin X. Boeckman,[a] Lauren Lessor,[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 an important human pathogen due to the wide range of infections it can cause and its emerging drug resistance. Isolation and characterization of phage infecting *K. pneumoniae* could be important for future therapeutic applications. Here, we report the complete genome sequence of the T4-like *Klebisella pneumoniae* myophage Mineola.

Klebsiella pneumoniae is an important opportunistic pathogen due to the continued emergence of highly drug-resistant strains (1), which carry the plasmid-borne and highly mobile *K. pneumoniae* carbapenemases (bla_{KPC}) (2). Isolation and characterization of phage infecting *K. pneumoniae* could be important for future therapeutic applications.

The myophage Mineola was isolated from activated sludge from the municipal wastewater in Bryan, TX, using a plasmid-cured derivative of a KPC-positive (KPC^+) *K. pneumoniae* clinical isolate of sequence type 258 as the host. Host bacteria were cultured on tryptic soy broth or agar (Difco) at 37°C with aeration. Phage were isolated and propagated with the soft agar overlay method (3). Phage genomic DNA was prepared using a modified Promega Wizard DNA cleanup kit protocol, as described previously (4). Pooled indexed DNA libraries were prepared with the Illumina TruSeq Nano DNA LT kit, and the sequence was obtained with the Illumina MiSeq platform with the MiSeq v2 500-cycle reagent kit, following the manufacturer's instructions; this produced 434,532 paired-end reads for the index containing the phage genome. FastQC 0.11.5 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), FASTX-Toolkit 0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/download.html), and SPAdes 3.5.0 (5) were used for read quality control, read trimming, and read assembly, respectively. The genome sequence was closed with PCR with primers (5′-GCCACCCATCATC AAACATATC-3′, 5′-CATCGGGTCGTCGTTCTAAA-3′) to face off the ends of the assembled contig and Sanger sequencing of the resulting product, and the contig sequence was manually corrected to match the resulting Sanger sequencing read. GLIMMER 3.0 (6) and MetaGeneAnnotator 1.0 (7) were used to predict protein-coding genes, which were then manually verified, and tRNA gene prediction was done with ARAGORN 2.36 (8). Putative protein functions were assigned based on sequence homology detected with BLASTp 2.2.28 (9), and conserved domains were detected with InterProScan 5.15-5.40 (10). All analyses were performed with default settings via the Center for Phage Technology (CPT) Galaxy (11) and WebApollo (12) interfaces (https://cpt.tamu.edu).

The phage Mineola genome sequence was assembled into 166,130 bp at 360.5-fold coverage. The genome contains 276 protein-coding genes and 16 tRNAs. Mineola has a GC content of 39.5%, which is 17.65% lower than that of its host (57.14%) (13). It shares 94.1% nucleotide similarity by progressiveMAUVE (version 2.4.0) (14) with *Klebsiella* phage JD18 (GenBank accession number KT239446). Mineola is T4 like, with 204 proteins that share homology with phage T4, as determined by BLASTp (E value ≤ 10^{-5}). The main divergences from T4 are in conserved hypothetical genes, which have

Citation Boeckman JX, Lessor L, Gill JJ, Liu M. 2019. Complete genome sequence of *Klebsiella pneumoniae* myophage Mineola. Microbiol Resour Announc 8:e00257-19. https://doi.org/10.1128/MRA.00257-19.

Editor Frank J. Stewart, Georgia Institute of Technology

Copyright © 2019 Boeckman et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

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

Received 4 March 2019
Accepted 3 April 2019
Published 25 April 2019

no homologs in T4 but are conserved among other *K. pneumoniae* phages, like phage JD18. The Mineola UvsW-like DNA helicase is encoded by two genes; this is a conserved feature among T4-like phages, and the single-gene UvsW in the T4 genome record (GenBank accession number NC_000866) is likely due to a sequencing error (15). All Mineola capsid and tail components have homologs in phase T4 except the putative distal subunit of the long tail fiber, which is 363 residues longer than its T4 gp37 counterpart and is more similar to the T5 L-shaped tail fiber (E value = 4^{-103}). A homing endonuclease that shares homology with T4 SegB is embedded in a region containing several tRNAs, which is thought to facilitate spreading of tRNA genes among T4-like phages (16).

Data availability. The genome sequence of phage Mineola was deposited under GenBank accession number MH333064. The associated BioProject, SRA, and BioSample accession numbers are PRJNA222858, SRR8788212, and SAMN11259693, respectively.

ACKNOWLEDGMENTS

This work was supported by funding from the National Science Foundation (awards EF-0949351 and DBI-1565146) and by the National Institutes of Health (NIAID award 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 the Department of Biochemistry and Biophysics.

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.

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. 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.
3. Adams MK. 1959. Bacteriophages. Interscience Publishers, New York, NY.
4. Summer EJ. 2009. Preparation of a phage DNA fragment library for whole genome shotgun sequencing, p 27–46. *In* Clokie MR, Kropinski AM (ed), Bacteriophages. Methods in molecular biology. Humana Press, Hoboken, NJ.
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 H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, 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 PJA, Grüning 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. Hoffmann M, Luo Y, Lafon PC, Timme R, Allard MW, McDermott PF, Brown EW, Zhao S. 2013. Genome sequences of *Salmonella enterica* serovar Heidelberg isolates isolated in the United States from a multistate outbreak of human Salmonella infections. Genome Announc 1:e00004-12. https://doi.org/10.1128/genomeA.00004-12.
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. Kerr ID, Sivakolundu S, Li Z, Buchsbaum JC, Knox LA, Kriwacki R, White SW. 2007. Crystallographic and NMR analyses of UvsW and UvsW.1 from bacteriophage T4. J Biol Chem 282:34392–34400. https://doi.org/10.1074/jbc.M705900200.
16. Brok-Volchanskaya VS, Kadyrov FA, Sivogrivov DE, Kolosov PM, Sokolov AS, Shlyapnikov MG, Kryukov VM, Granovsky IE. 2008. Phage T4 SegB protein is a homing endonuclease required for the preferred inheritance of T4 tRNA gene region occurring in co-infection with a related phage. Nucleic Acids Res 36:2094–2105. https://doi.org/10.1093/nar/gkn053.