Complete Genome Sequence of *Klebsiella aerogenes* Myophage Metamorpho

Kyra E. Groover,[a] Kameron Garza,[a] James Clark,[a] Isla Hernandez,[a] Jason J. Gill,[a] Mei Liu[a]

aCenter for Phage Technology, Texas A&M University, College Station, Texas, USA

ABSTRACT The bacterium *Klebsiella aerogenes* is an opportunistic pathogen that often infects hospitalized patients and those who are immunocompromised. *K. aerogenes* in some cases can become resistant to antibiotic treatment. Being a potential therapeutic, Metamorpho is a T4-like myophage targeting *K. aerogenes*.

Klebsiella aerogenes is an opportunistic pathogen that is generally associated with urinary tract infections, abdominal infections, septicemia, and pneumonia, with a majority of infections associated with this pathogen being hospital acquired (1). There is a high incidence of resistance to β-lactam antibiotics in *K. aerogenes* (1). Phage Metamorpho is able to lyse and kill *K. aerogenes*, and its annotated genome is described here.

Metamorpho was isolated from wastewater collected from the wastewater plant in Madisonville, Texas, using *K. aerogenes* strain ATCC 13048 with the soft-agar overlay method (2). The phage was propagated at 37°C on LB agar seeded with the host bacterial strain. The phage morphology was visualized via transmission electron microscopy (TEM) of virions stained with 2% (wt/vol) uranyl acetate (3) at the Texas A&M Microscopy and Imaging Center. Phage DNA was extracted using the Promega Wizard DNA extraction system following the steps described previously (4) and prepared for sequencing using the Illumina Nextera kit with 300-bp inserts. The prepared DNA was subjected to sequencing with a MiSeq system using 300-cycle v2 chemistry. The sequences returned were checked for quality by FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc). The library index containing this phage generated 686,406 sequencing reads, and ~5% of the total reads were assembled into SPAdes v3.5.0 (3) into a contig corresponding to this phage with 10× coverage. The genome was closed using PCR, with the primers 5′-GCTATTCTATCCCAACGGTCAG-3′ and 5′-GAATAGGAT CAACCGAGTTACCG-3′, and Sanger sequencing of the PCR product. Structural annotation was conducted with GLIMMER v3 (5) and MetaGeneAnnotator v1.0 (6) for identification of putative genes. ARAGORN v2.36 (7) was used to identify tRNA genes. Protein functions were predicted by InterProScan v5.33 (8), BLAST v2.9.0 (9) with the NCBI non-redundant and Swiss-Prot databases (10), TMHMM v2.0 (11), LipoP (12), and HHpred (13) with default settings. The whole-genome sequence was compared to those of other phages via progressiveMauve v2.4 (14). All of the tools used for analysis were run on the Center for Phage Technology (CPT) Galaxy and Apollo interfaces (15–17) with default settings (https://cpt.tamu.edu/galaxy-pub).

The complete genome length of phage Metamorpho is 171,475 bp. It has a lower GC content (38.5%) than its host (54.1%) (18). A total of 287 protein-coding genes and 15 tRNAs were predicted, for a coding density of 94.0%. Metamorpho is most closely related to other T4-like *Klebsiella* phages; the most notable are JD18 (GenBank accession number KT239446), Mineola (MH333064), and KP1 (MG751100). Metamorpho is also related to the canonical *Escherichia coli* phage T4, with 75.7% DNA sequence similarity to T4 by BLASTn analysis and 177 proteins shared with T4 by BLASTp (E <10^{-5}).

Not surprisingly, Metamorpho showed T4-like myophage TEM morphology (image not shown). Metamorpho was predicted by PhageTerm (19) to package its DNA by headful packaging, and the genome was reopened at the *rII* locus to retain synteny with other T4-like phages. Metamorpho gp163 is a homolog of T4 protein e.6, which is conserved in a number of T4-like phages. Analysis of gp163 with HHpred indicated a high-quality alignment (99.9%) with *E. coli* phage shock protein PspA (4WHE_A), which is associated with the cellular stress response (20).

Data availability. The genome sequence of phage Metamorpho was deposited under GenBank accession number MT701588 and BioSample accession number SAMN14609635. The BioProject accession number is PRJNA222858, and the SRA accession number is SRR11558337.

REFERENCES

1. Davin-Regli A, Lavigne JP, Pagès JM. 2019. *Enterobacter* spp.: update on taxonomy, clinical aspects, and emerging antimicrobial resistance. Clin Microbiol Rev 32:e00002-19. https://doi.org/10.1128/CMR.00002-19.
2. Adams MK. 1959. Bacteriophages. Interscience Publishers, Inc., New York, NY.
3. Valentine RC, Shapiro BM, Stadtman ER. 1968. Regulation of glutamine synthetase. XII. Electron microscopy of the enzyme from *Escherichia coli*. Biochemistry 7:2143–2152. https://doi.org/10.1021/bi00846a017.
4. Summer EJ. 2009. Preparation of a phage DNA fragment library for whole genome shotgun sequencing. Methods Mol Biol 502:27–46. https://doi.org/10.1007/978-1-60327-565-1_4.
5. 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.
6. 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.
7. 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.
8. 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.
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. The UniProt Consortium. 2018. UniProt: the universal protein knowledgebase. Nucleic Acids Res 46:2699. https://doi.org/10.1093/nar/gky092.
11. 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.
12. 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.
13. 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.
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. Ramsey J, Rasche H, Maughmer C, Criscione A, Mijalis E, Liu M, Hu JC, Young R, Gill JJ. 2020. Galaxy and Apollo as a biologist-friendly interface for high-quality cooperative phage genome annotation. PLoS Comput Biol 16:e1008214. https://doi.org/10.1371/journal.pcbi.1008214.
16. Dunn NA, Unni DR, Diesh C, Munoz-Torres M, Harris NL, Yao E, Rasche H, Holmes IH, Elsik CG, Lewis SE. 2019. Apollo: democratizing genome annotation. PLoS Comput Biol 15:e1006790. https://doi.org/10.1371/journal.pcbi.1006790.
17. Jalili V, Afgan E, Gu Q, Clements D, Blankenberg D, Goecks J, Taylor J, Nekrutenko A. 2020. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. Nucleic Acids Res 48:W395–W402. https://doi.org/10.1093/nar/gkaa434.
18. Shin SH, Kim S, Kim JY, Lee S, Um Y, Oh MK, Kim YR, Lee J, Yang KS. 2012. Complete genome sequence of *Enterobacter aerogenes* KCTC 2190. J Bacteriol 194:2373–2374. https://doi.org/10.1128/JB.00028-12.
19. 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.
20. Flores-Kim J, Darwin AJ. 2016. The phage shock protein response. Annu Rev Microbiol 70:83–101. https://doi.org/10.1146/annurev-micro-102215-095359.