Microbiology®
Resource Announcements

Complete Genome Sequence of *Alcaligenes faecalis* Phage Piluca

Joshua Wan,ᵃ Zihao Yu,ᵃ James Clark,ᵃ,ᵇ Tram Le,ᵃ,ᵇ Ryland Young,ᵃ,ᵇ Mei Liuᵃ,ᵇ

ᵃDepartment of Biochemistry and Biophysics, Texas A&M University, College Station, Texas, USA
ᵇCenter for Phage Technology, Texas A&M University, College Station, Texas, USA

ABSTRACT *Alcaligenes faecalis* is an opportunistic pathogen exhibiting drug resistance. Here, the 35,451-bp genome of *A. faecalis* phage Piluca is described. Piluca is not closely related to any isolated phages in the NCBI database. Piluca possesses genes encoding CI-like and Cro-like repressors and a tyrosine integrase, suggesting its temperate lifestyle.

*A*lcaligenes faecalis is a Gram-negative bacterium commonly found in watery environments and causes opportunistic infections in humans that are difficult to treat because of increased resistance to antibiotics (1, 2). Phages are of interest as potential therapeutic agents to control infections associated with this bacterium.

Phage Piluca was isolated in 2019 from a wastewater sample collected in a wastewater treatment plant in Madisonville, TX, using *A. faecalis* ATCC strain 8750 as the isolation host. The bacterial host strain was cultured using tryptic soy broth or agar (Difco) at 37°C with aeration, and the phage was isolated and propagated using the soft agar overlay method (3). Phage DNA was purified from ~8 mL phage lysate using the Promega Wizard DNA cleanup system as previously described (4). DNA sequencing libraries were prepared as 300-bp inserts using a Swift 2S Turbo kit and sequenced on an Illumina MiSeq instrument with paired-end 150-bp reads using V2 300-cycle chemistry. This produced 203,894 total raw reads, which were then quality controlled using FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc) and trimmed with FastX-Toolkit v0.11.6 (http://hannonlab.cshl.edu/fastx_toolkit/). The raw reads were assembled into a single contig using SPAdes v3.5.0 (5) with 244-fold coverage, and the contig sequence was completed by PCR amplifying off the contig ends (primers: forward 5′-CGGCAATGGGAATGCATGAG-3′, reverse 5′-TTGGTGAAA GAATGGCCCGT-′3) and Sanger sequencing of the resulting product followed by manual curation. The genome was annotated using the CPT Galaxy-Apollo phage annotation platform (https://cpt.tamu.edu/galaxy-pub) (6–8). Structural annotation was performed with GLIMMER v3 (9) and MetaGeneAnnotator v1.0 (10). tRNAs were detected using ARAGORN v2.36 (11) and tRNAscan-SE v2.0 (12). Gene functions were predicted with InterProScan v5.48 (13), BLAST v2.9.0 (14), TMHMM v2.0 (15), HHPred (16), LipoP v1.0 (17), and SignalP v5.0 (18). BLAST searches were compared to the NCBI nonredundant (nr) and SwissProt databases (19). Genome-wide DNA sequence similarity was calculated with progressive Mauve v2.4 (20). All tools were run at default settings.

Phage Piluca has a 35,451-bp genome with 41% G+C content and 96.1% coding density. Piluca is not closely related enough to any isolated phages in the NCBI nr database to be placed in a phage type group in the *Podoviridae* family, as it is lacking genome-wide nucleotide identity and shares no more than 12 proteins with any other isolated phage (BLASTp E value, <0.001). Piluca is predicted to be a podophage based on its tail proteins that show significant similarity (BLASTp E value, <0.001) to that carried by *Enterobacteria* podophage T7 (GenBank accession number NC_001604) and phage epsilon15 (NC_004775). Piluca possesses genes encoding CI-like and Cro-like

transcriptional regulators and a tyrosine integrase, suggesting it possesses a temperate lifestyle. Piluca is related on the nucleotide level to prophage elements identified in multiple *Acinetobacter baumannii* strains, such as *A. baumannii* strain VB82 (CP050385, ~65% overall similarity based on BLASTn). Overall, 23 of the total 49 predicted protein-coding genes could be assigned a putative function. Piluca has genes encoding a class I holin and a signal-arrest-release (SAR) endolysin, but no spanins were identified. No tRNAs were identified in the genome. PhageTerm (21) analysis of the raw sequencing reads indicated that Piluca packages its DNA by a headfull mechanism.

Data availability. The Phage Piluca genome sequence was deposited in GenBank under accession number MZ326864. The associated BioProject, SRA, and BioSample accession numbers are PRJNA222858, SRR14095246, and SAMN18509619, respectively.

REFERENCES

1. Bizet J, Bizet C. 1997. Strains of *Alcaligenes faecalis* from clinical material. J Infect 35:167–169. https://doi.org/10.1016/s0163-4453(97)91710-2.
2. Huang C. 2020. Extensively drug-resistant *Alcaligenes faecalis* infection. BMC Infect Dis 20:833. https://doi.org/10.1186/s12879-020-05557-8.
3. Clokie MR, Kropinski AM, Lavigne R. 2009. Bacteriophages. Springer, Berlin, Germany.
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. 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. 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.
7. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Gruning BA, Guerler A, Hillman-Jackson J, Hiltemann S, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res 46: W537–W544. https://doi.org/10.1093/nar/gky379.
8. 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.
9. 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.
10. 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.
11. 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.
12. Chan PP, Lowe TM. 2019. tRNAscan-SE: searching for tRNA genes in genomic sequences. Methods Mol Biol 1962:1–14. https://doi.org/10.1007/978-1-4939-9173-0_1.
13. 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.
14. 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.
15. 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.
16. 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.
17. 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.
18. Almagro Armenteros JJ, Tsirigos KD, Sonderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol 37: 420–423. https://doi.org/10.1038/s41587-019-0036-z.
19. UniProt Consortium T. 2018. UniProt: the universal protein knowledgebase. Nucleic Acids Res 46:2699. https://doi.org/10.1093/nar/gky092.
20. 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.
21. 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.