



Complete Genome Sequence of *Stenotrophomonas maltophilia* Siphophage Siara

John Marmion,^a Nathaniel Tate,^a James Clark,^{a,b} Tram Le,^{a,b} Ben Burrowes,^c  Mei Liu^{a,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

^cBB Phage Consultancy, LLC, Georgetown, Texas, USA

ABSTRACT *Stenotrophomonas maltophilia* is associated with an increasing incidence of nosocomial infections. Here, we describe the isolation and genome annotation of *S. maltophilia* siphophage Siara. Its 61,427-bp genome is currently related only to one phage in the NCBI database, namely, *S. maltophilia* phage Salva, and is not related to any prophages.

The increasing incidence of nosocomial infection caused by *Stenotrophomonas maltophilia* is concerning due to some strains being multidrug resistant and the significant infection fatality-to-case ratio (1). The isolation and genome annotation of phage Siara, a potential therapeutic agent for controlling multidrug-resistant *S. maltophilia* are described here.

Phage Siara was isolated in 2019 from an influent water sample collected from a wastewater treatment plant in Beaumont, TX (Global Positioning System [GPS] coordinates of 30.20078, -94.10807), using *S. maltophilia* (ATCC 17807) as the host. The host strain was propagated aerobically at 30°C in tryptone nutrient (0.5% tryptone, 0.25% yeast extract, 0.1% glucose, and 0.85% NaCl [wt/vol]) broth or agar. Phage isolation and propagation were done using the soft agar overlay method (2). Phage DNA was purified from polyethylene glycol (PEG)-precipitated phage particles from ~8 mL phage lysate (>10⁹ PFU/mL) using a Promega Wizard DNA cleanup system following the manufacturer's protocol (3). The purified DNA was prepared as 300-bp inserts using a Swift 2S Turbo library preparation kit and sequenced on an Illumina MiSeq instrument with paired-end 150-bp reads using V2 300-cycle chemistry. A total of 174,160 raw sequence reads were quality controlled with FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>) and trimmed using the FASTX-Toolkit v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/) to generate a total of 125,775 trimmed reads, from which a single contig with 129-fold coverage was assembled with SPAdes v3.5.0 (4). Contig end sequences were verified by PCR and Sanger sequencing the resulting product using primers 5'-TGCTGCCGTTACAAAACAG-3' and 5'-TCCTGACTCTACCCACCCTG-3'. The phage termini were predicted using PhageTerm (5). The structural annotation was done using GLIMMER v3 (6) and MetaGeneAnnotator v1.0 (7), and tRNAs were detected with ARAGORN v2.38 (8) and tRNAScan-SE v2.0 (9). The gene functions were predicted using InterProScan v5.48 (10), TMHMM v2.0 (11), HHpred (12), LipoP v1.0 (13), and SignalP v5.0 (14), as well as BLAST (15) searches against the following databases: NCBI nonredundant and SwissProt (16). Rho-independent termination sites were annotated using TransTermHP v2.09 (17). progressiveMauve v2.4 (18) was used to calculate the genome-wide similarity. Analysis tools were accessed through the Center for Phage Technology (CPT) Galaxy-Apollo platform (<https://cpt.tamu.edu/galaxy-pub>) (19–21), and all software were used at their default settings. The morphology of phage Siara was determined by transmission electron microscopy (TEM) and viewing samples stained negatively with 2% (wt/vol) uranyl acetate at the Texas A&M Microscopy and Imaging Center.

Editor Catherine Putonti, Loyola University Chicago

Copyright © 2022 Marmion et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

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

The authors declare no conflict of interest.

Received 22 February 2022

Accepted 19 April 2022

Published 2 May 2022

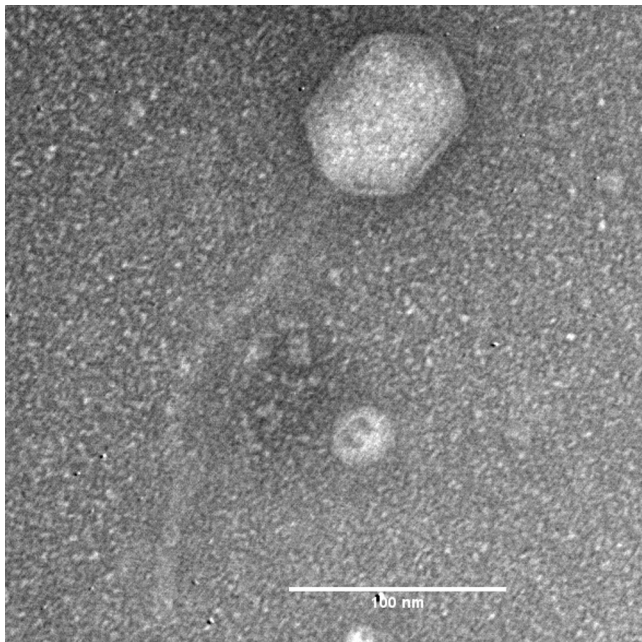


FIG 1 Transmission electron micrograph of phage Siara. Phage particles were diluted with TEM buffer (20 mM NaCl, 10 mM Tris-HCl [pH 7.5], and 2 mM MgSO₄) and captured on a freshly glow-discharged, Formvar carbon-coated grid. The grids were stained with 2% (wt/vol) uranyl acetate and observed on a JEOL 1200 EX transmission electron microscope at a 100-kV accelerating voltage at the Microscopy and Imaging Center at Texas A&M University.

Phage Siara is determined to have a siphophage morphology (Fig. 1). Siara has a complete genome length of 61,427 bp, a coding density of 93.9%, and a G+C content of 56.5%. A total of 100 protein-coding genes and 3 tRNA genes were identified in the Siara genome. At the time of writing of this manuscript, phage Siara is closely related to only one phage deposited in GenBank, namely, phage Salva (GenBank accession number [MW393850](https://doi.org/10.1093/genbank/MW393850)) (22), sharing 86.7% nucleotide identity over 73% Siara genome coverage as determined by BLASTn. Similar to the long terminal repeat region (3,973 bp) identified in phage Salva, PhageTerm analysis identified a putative 2,528-bp terminal repeat region in the Siara genome. Phage Siara and Salva are currently unclassified members within the *Siphoviridae* family. Based on BLASTn at an overall nucleotide similarity cut off 30% (calculated by using the percent identity times the percent aligned length, as determined by BLASTn), Siara is not closely related to any prophage elements in NCBI bacterial genomes.

Data availability. Siara was deposited in GenBank with accession number [MZ326859](https://doi.org/10.1093/genbank/MZ326859). The associated BioProject, SRA, and BioSample accession numbers are [PRJNA222858](https://doi.org/10.1093/bioinformatics/PRJNA222858), [SRR14095253](https://doi.org/10.1093/bioinformatics/SRR14095253), and [SAMN18509351](https://doi.org/10.1093/bioinformatics/SAMN18509351), respectively.

ACKNOWLEDGMENTS

Funding was provided by the National Science Foundation (awards EF-0949351 and DBI-1565146) and from the Center for Phage Technology (CPT).

Carlos Gonzalez, CPT, provided the *Stenotrophomonas* strain.

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

REFERENCES

1. Brooke JS. 2012. *Stenotrophomonas maltophilia*: an emerging global opportunistic pathogen. *Clin Microbiol Rev* 25:2–41. <https://doi.org/10.1128/CMR.00019-11>.
2. Adams MH. 1959. *Bacteriophages*. Interscience Publishers, New York, NY.
3. 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.
4. 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>.
5. Garneau JA-O, Depardieu F, Fortier LC, Bikard DA-O, Monot MA-O. 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>.
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. 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.
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, Scheremetjov 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. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. 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. 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>.
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. 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>.
15. 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>.
16. The UniProt Consortium. 2018. UniProt: the universal protein knowledge-base. *Nucleic Acids Res* 46:2699. <https://doi.org/10.1093/nar/gky092>.
17. Kingsford CL, Ayanbule K, Salzberg SL. 2007. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol* 8:R22. <https://doi.org/10.1186/gb-2007-8-2-r22>.
18. 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>.
19. 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>.
20. 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, Hiltmann 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>.
21. Dunn NA, Unni DR, Diesh C, Munoz-Torres M, Harris NL, Yao E, Rasche H, Holmes IH, Elisk CG, Lewis SE. 2019. Apollo: democratizing genome annotation. *PLoS Comput Biol* 15:e1006790. <https://doi.org/10.1371/journal.pcbi.1006790>.
22. Jefferson B, Yao G, Clark J, Le T, Gonzalez CF, Liu M, Burrows B. 2021. Complete genome sSequence of *Stenotrophomonas maltophilia* siphophage Salva. *Microbiol Resour Announc* 10:e00083-21. <https://doi.org/10.1128/MRA.00083-21>.