



# Complete Genome Sequence of *Stenotrophomonas maltophilia* Podophage Piffle

Megan Kirchhoff,<sup>a</sup> Carlos Ortega,<sup>a</sup> James Clark,<sup>a,b</sup> Tram Le,<sup>a,b</sup> Ben Burrowes,<sup>c</sup>  Mei Liu<sup>a,b</sup>

<sup>a</sup>Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas, USA

<sup>b</sup>Center for Phage Technology, Texas A&M University, College Station, Texas, USA

<sup>c</sup>BB Phage Consultancy, LLC, Georgetown, Texas, USA

**ABSTRACT** *Stenotrophomonas maltophilia* is an emerging multidrug-resistant opportunistic human pathogen causing various nosocomial infections. Here, we characterize the genome of *S. maltophilia* podophage Piffle. Its 76,332-bp genome is most closely related to the N4-like *S. maltophilia* podophage Pokken, with over 86% genome-wide nucleotide identity and 84 shared proteins.

*Stenotrophomonas maltophilia* is an emerging multidrug-resistant (MDR) opportunistic pathogen causing various nosocomial infections, and multiple virulence traits make it especially harmful to immunocompromised patients (1). Phage therapy has the potential to treat MDR infections (2), and we describe here the genome of Piffle, a potential therapeutic phage for *S. maltophilia* infections.

Piffle was isolated in September 2019 from a trickle filter effluent sample collected from a wastewater treatment plant in Beaumont, TX, using *S. maltophilia* (ATCC 18707) as the propagation host. Host bacteria were grown on tryptone nutrient (0.5% tryptone, 0.25% yeast extract, 0.1% glucose, 0.85% NaCl [wt/vol]) broth or agar at 30°C, and phages were propagated by the soft-agar overlay method (3). Phage morphology was determined via negative staining with 2% (wt/vol) uranyl acetate (4), and imaging was done by transmission electron microscopy at the Texas A&M Microscopy and Imaging Center. Phage genomic DNA was purified from ~8 mL phage lysate using the Promega Wizard DNA cleanup system as previously described (5), 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 140,904 raw sequence reads were quality controlled with FastQC ([www.bioinformatics.babraham.ac.uk/projects/fastqc](http://www.bioinformatics.babraham.ac.uk/projects/fastqc)) and trimmed with the FASTX-Toolkit v0.0.14 ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)). Then, 67,758 trimmed reads were used for assembly with SPAdes v3.5.0 (6). A raw contig with 68-fold coverage was assembled. Since the raw contig assembled by SPAdes is usually opened at a random spot in the middle of the genome, and the contig ends often have redundant or missing bases, PCR amplification on the genomic DNA using primers designed off the contig ends (forward, 5'-GAGTAGCGAGC CATGACGAA-3' and reverse, 5'-ACTGAGGTCGAGGTCGAGAA-3' for phage Piffle) followed by Sanger sequencing of the PCR product allows us to verify sequences in that region. Analyses were done using the CPT Galaxy-Apollo platform (<https://cpt.tamu.edu/galaxy-pub>) (7–9). PhageTerm was used to predict genomic termini (10). Structural annotation was performed to identify protein-coding genes using GLIMMER v3 and MetaGeneAnnotator v1.0, to identify tRNAs using ARAGORN v2.36 and tRNAscan-SE v2.0, and to identify rho-independent terminators using TransTermHP v2.09 (11–15). Functions were assigned to genes based on conserved protein domains predicted by InterProScan v5.48 and BLAST v2.9.0 against the NCBI nonredundant and SwissProt databases (16–18), with a maximum expectation value cutoff of 0.001. Further

**Editor** Kenneth M. Stedman, Portland State University

**Copyright** © 2022 Kirchhoff 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/).

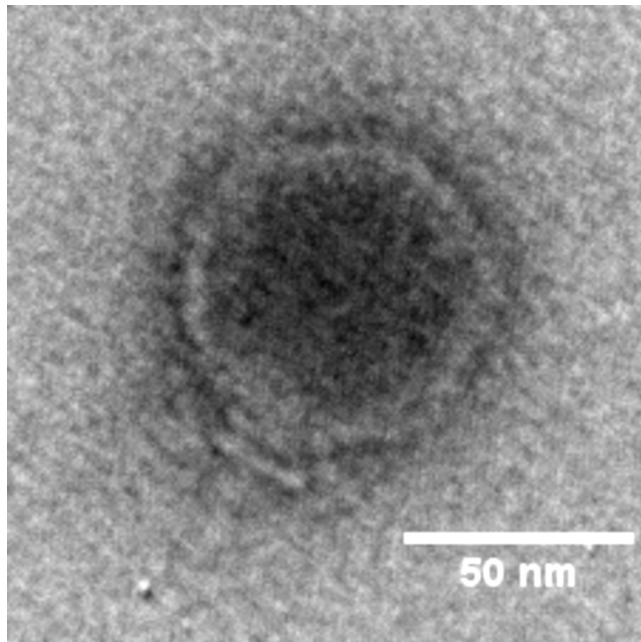
Address correspondence to Mei Liu, [meiliu@tamu.edu](mailto:meiliu@tamu.edu).

The authors declare no conflict of interest.

**Received** 20 February 2022

**Accepted** 10 March 2022

**Published** 23 March 2022



**FIG 1** Transmission electron micrograph (TEM) of phage Piffle. Phage particles were negatively stained with 2% (wt/vol) uranyl acetate and observed on a JEOL 1200 EX TEM at 100 kV accelerating voltage at the Microscopy and Imaging Center at Texas A&M University.

analyses include determining transmembrane domains with TMHMM v2.0, potential span-ins with LipoP v1.0, and signal peptides with SignalP v5.0 (19–21). External analyses were performed with the HHpred, UniProtKB, SwissProt, NCBI Conserved Domains v3.18, and TIGRFAMs v15.0 reference databases (18, 22–25). progressiveMauve v2.4 was used to calculate genome-wide DNA sequence similarities (26). All software was used at default settings.

Phage Piffle has a podophage morphology (Fig. 1). The 76,332-bp genome of Piffle has a coding density of 91.8% and a G+C content of 54.9%, which is lower than the 66.8% G+C content of its host (27). A total of 6 tRNA genes and 90 protein-coding genes were predicted, 38 of which were assigned functions. Piffle is most closely related to the N4-like phage of *S. maltophilia*, Pokken, with over 86% genome-wide nucleotide identity determined by progressiveMauve and 84 shared proteins (BLASTp E value, <0.001). PhageTerm could not determine the genomic termini of Piffle even though it is predicted to possess terminal repeats based on its relatedness to Pokken, whereas 627-bp direct terminal repeats were predicted for Pokken, as expected for N4-like phages (28). Phage Piffle is also related to another N4-like phage, *Xanthomonas* phage RiverRider (GenBank accession number [NC\\_048703](https://ncbi.nlm.nih.gov/nuccore/NC_048703)) (29), sharing approximately 53% overall nucleotide identity as determined by progressiveMauve.

**Data availability.** The Piffle genome was deposited in GenBank with accession number [MZ326857](https://ncbi.nlm.nih.gov/nuccore/MZ326857). The associated BioProject, SRA, and BioSample accession numbers are [PRJNA222858](https://ncbi.nlm.nih.gov/bioproject/PRJNA222858), [SRR14095255](https://ncbi.nlm.nih.gov/sra/SRR14095255), and [SAMN18509294](https://ncbi.nlm.nih.gov/biosample/SAMN18509294), 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. This announcement 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. Adegoke AA, Stenström TA, Okoh AI. 2017. *Stenotrophomonas maltophilia* as an emerging ubiquitous pathogen: looking beyond contemporary antibiotic therapy. Front Microbiol 8:2276. <https://doi.org/10.3389/fmicb.2017.02276>.

3. Adams MH. 1956. Bacteriophages. Interscience Publishers, Inc., New York, NY.
4. 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>.
5. 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](https://doi.org/10.1007/978-1-60327-565-1_4).
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. 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>.
8. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Grüning 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>.
9. 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>.
10. Garneau JR, Depardieu F, Fortier L-C, 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>.
11. 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>.
12. 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>.
13. 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>.
14. 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](https://doi.org/10.1007/978-1-4939-9173-0_1).
15. 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>.
16. 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>.
17. 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>.
18. The UniProt Consortium. 2018. UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 46:2699. <https://doi.org/10.1093/nar/gky092>.
19. 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>.
20. 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>.
21. Almagro Armenteros JJ, Tsirigos KD, Sønderby 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>.
22. Gabler F, Nam SZ, Till S, Mirdita M, Steinegger M, Söding J, Lupas AN, Alva V. 2020. Protein sequence analysis using the MPI bioinformatics toolkit. *Curr Protoc Bioinformatics* 72:e108. <https://doi.org/10.1002/cpbi.108>.
23. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. *Nucleic Acids Res* 28:235–242. <https://doi.org/10.1093/nar/28.1.235>.
24. Yang M, Derbyshire MK, Yamashita RA, Marchler-Bauer A. 2020. NCBI's Conserved Domain Database and tools for protein domain analysis. *Curr Protoc Bioinformatics* 69:e90. <https://doi.org/10.1002/cpbi.90>.
25. Haft DH, Selengut JD, Richter RA, Harkins D, Basu MK, Beck E. 2013. TIGRFAMs and genome properties in 2013. *Nucleic Acids Res* 41:D387–D395. <https://doi.org/10.1093/nar/gks1234>.
26. 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>.
27. Lira F, Berg G, Martinez JL. 2017. Double-face meets the bacterial world: the opportunistic pathogen *Stenotrophomonas maltophilia*. *Front Microbiol* 8:2190. <https://doi.org/10.3389/fmicb.2017.02190>.
28. Hayden A, Martinez N, Moreland R, Liu M, Gonzalez CF, Gill JJ, Ramsey J. 2019. Complete genome sequence of *Stenotrophomonas* phage Pokken. *Microbiol Resour Announc* 8:e01095-19. <https://doi.org/10.1128/MRA.01095-19>.
29. Miller M, Deulio A, Holland C, Douthitt C, McMahon J, Wiersma-Koch H, Turechek WW, D'Elia T. 2020. Complete genome sequence of *Xanthomonas* phage RiverRider, a novel N4-like bacteriophage that infects the strawberry pathogen *Xanthomonas fragariae*. *Arch Virol* 165:1481–1484. <https://doi.org/10.1007/s00705-020-04614-6>.