Complete Genome Sequence of *Stenotrophomonas maltophilia* Siphophage Summit

Alexandra Bishop,[a] Xing-Han Zhang,[a] James Clark,[b] Tram Le,[b] Ben Burrowes,[b] Mei Liu[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 *Stenotrophomonas maltophilia* is an opportunistic pathogen exhibiting resistance to multiple antimicrobials. This study reports the complete genome of an *S. maltophilia* siphophage, Summit. Its genome of 95,728 bp has 148 protein-coding genes and 5 tRNAs, including 1 predicted suppressor tRNA. Possible target genes for the suppressor tRNA are not identified.

S tenotrophomonas maltophilia, a Gram-negative bacterium, is both a human opportunistic pathogen and potential biocontrol agent for use in bioremediation (1, 2). *S. maltophilia* exhibits intrinsic resistance to various classes of antimicrobials, causing the need for new drugs, such as bacteriophage therapy, to counter *S. maltophilia* infections (2, 3). This report summarizes the complete genome annotation of a novel *S. maltophilia* phage siphophage, Summit.

Bacteriophage Summit was isolated from a weaning foal swab sample provided by the Texas A&M Veterinary Clinic, College Station, TX. The phage was isolated and cultured by the soft agar overlay method (4) with *S. maltophilia* (ATCC 17807) as a propagation host grown aerobically at 30°C in tryptone nutrient broth. Following purification, phage DNA was isolated as previously described (5), prepared as 300-bp inserts using a Swift 2S Turbo kit, and sequenced on an Illumina MiSeq with pair-end 150-bp reads using a V2 300-cycle chemistry. Using FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc) and FASTX-Toolkit v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/) resulted in 129,787 trimmed reads. Summit's genome was then assembled by SPAdes v3.5.0 into a single contig with 88-fold coverage (6). Forward (5′-TGGAAGAGAAGGCCACGAAC-3′) and reverse (5′-CCGAGTCGAGGTAGAACGTG-3′) closure primers were used to close the genome by PCR and Sanger sequencing. The genome was annotated using the CPT Galaxy-Apollo phage annotation platform (7–9). Structural annotation was completed using Glimmer v3 (10) and MetaGeneAnnotator v1.0 (11). tRNAs were identified with ARAGORN v2.36 (12) and tRNAScan-SE v2.0 (13). Functional gene annotation was completed using InterProScan v5.48 (14), BLAST v2.9.0 (15), TMHMM v2.0 (16), HHPred (17), LipoP v1.0 (18), and SignalP v5.0 (19). BLAST comparison was run against two databases, NCBI nr and SwissProt (20). Genome-wide DNA sequence similarities were calculated by ProgressiveMauve v2.4 (21). All analyses were done at default settings. Phage morphology was determined by staining samples with 2% (wt/vol) uranyl acetate (22) and viewing the samples by transmission electron microscopy (TEM) at the Texas A&M Microscopy and Imaging Center.

Phage Summit has a siphophage morphology (Fig. 1). It has a genome of 95,728 bp with a 58.5% GC content and 93.7% coding density. Annotation of the genome revealed 148 protein-coding genes and 5 tRNAs, including 1 predicted suppressor tRNA. Analysis was done to identify possible target genes for the predicted amber stop codon suppressor tRNA. No significant genes were identified that would require the suppressor tRNA or that were lacking a Shine-Dalgarno sequence. There are 12 genes in the genome that have amber stop codons. None of those genes are followed by a
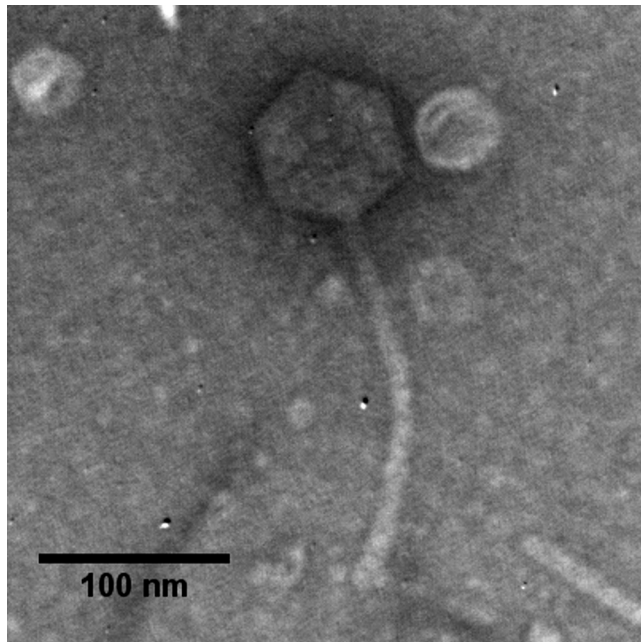
FIG 1 Transmission electron micrograph (TEM) of phage Summit. Phage particles were diluted with TEM buffer (20 mM NaCl, 10 mM Tris-HCl [pH 7.5], 2 mM $MgSO_4$) and captured on freshly glow-discharged, Formvar carbon-coated grids. The grids were stained with 2% (wt/vol) uranyl acetate and observed on a JEOL 1200 EX TEM at 100-kV accelerating voltage at the Microscopy and Imaging Center at Texas A&M University.

predicted gene that was in frame that could be fused to its upstream neighbor by amber suppression, and no identifiable conserved domains or BLAST hits were found in downstream sequences that would be the product of such a readthrough. No tRNA synthetase protein was predicted during functional annotation. Therefore, the purpose of the predicted suppressor tRNA remains unknown. At the time of the analysis, *Stenotrophomonas* phage vB_SmaS_DLP_5 (GenBank accession number NC_042082.1) was the only known phage closely related to Summit, sharing 86.4% nucleotide similarity as determined by ProgressiveMauve and 139 proteins (BLASTp; E value < 0.001).

Data availability. Summit was deposited in GenBank with accession number MZ326862. The associated BioProject, SRA, and BioSample accession numbers are PRJNA222858, SRR14095250, and SAMN18509476, respectively.

REFERENCES

1. An S-Q, Berg G. 2018. *Stenotrophomonas maltophilia*. Trends Microbiol 26: 637–638. https://doi.org/10.1016/j.tim.2018.04.006.
2. Brooke JS, Di Bonaventura G, Berg G, Martinez J-L. 2017. A multidisciplinary look at *Stenotrophomonas maltophilia*: an emerging multi-drug-resistant global opportunistic pathogen. Front Microbiol 8:1511. https://doi.org/10.3389/fmicb.2017.01511.
3. Denton M, Kerr KG. 1998. Microbiological and clinical aspects of infection associated with *Stenotrophomonas maltophilia*. Clin Microbiol Rev 11: 57–80. https://doi.org/10.1128/CMR.11.1.57.
4. Adams M. 1959. Bacteriophages. Interscience Publishers, Inc, New York, NY.
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.
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. 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. 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.

10. 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.

11. 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.

12. 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.

13. 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.

14. 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.

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. 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.

17. 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.

18. 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.

19. 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.

20. UniProt Consortium. 2018. UniProt: the universal protein knowledgebase. Nucleic Acids Res 46:2699. https://doi.org/10.1093/nar/gky092.

21. 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.

22. 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.