AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Proteus mirabilis* Siphophage Saba

James Nguyen,[a] Laith Harb,[a] Russell Moreland,[a] Mei Liu,[a] Jason J. Gill,[a] Jolene Ramsey[a]

[a]Center for Phage Technology, Texas A&M University, College Station, Texas, USA

ABSTRACT *Proteus mirabilis* is a Gram-negative enteric bacterium associated with complicated human urinary tract infections. Here, we present the complete genome annotation for *P. mirabilis* siphophage Saba. With a 60,056-bp genome and 75 predicted genes, Saba is most similar at the nucleotide and protein levels to phage Chi and Chi-like viruses.

Proteus mirabilis is a Gram-negative bacterium commonly found in soil, stagnant water, and sewage and is associated with various animal gastrointestinal tracts (1). *P. mirabilis*, known for its highly motile swarming lifestyle and formation of biofilms on catheters, is a leading cause of complicated urinary tract infections in humans (2). Bacteriophage applications to mitigate *P. mirabilis* disease are being explored (3), which led us to isolate and annotate *P. mirabilis* phage Saba.

The source for Saba was filtered (0.2-μm-pore-size filter) wastewater in College Station, TX. Host *P. mirabilis* strain ATCC 29906 was cultured aerobically at 37°C in nutrient broth/agar (BD). Phages were isolated by the soft-agar overlay method of Adams (4), and small phages were selected by the ability to pass through an Amicon Ultra-15 spin filter with a nominal 30-kDa molecular weight limit (Millipore-Sigma). To determine phage morphology, samples were negatively stained with 2% (wt/vol) uranyl acetate and imaged by transmission electron microscopy at the Texas A&M Microscopy and Imaging Center (5). Saba genomic DNA was purified according to Summer's modified shotgun library protocol (6). The genome was prepared as Illumina TruSeq libraries with the Nano low-throughput kit and sequenced on an Illumina MiSeq platform using V2 500-cycle chemistry for 250-bp paired-end reads. The 1,421,963 total reads in the phage-containing index were quality controlled using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). FASTX-Toolkit v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/) was used for read trimming, and then assembly into a single contig at 44.6-fold coverage was accomplished with SPAdes v3.5.0, with default parameters (7). Contig completion was confirmed by PCR off the ends (forward primer, 5′-TTTACCTTGGAGCACTTGATCG-3′, and reverse primer, 5′-GTAGGGCGGTATTGCGTTTAT-3′) and Sanger sequencing. PhageTerm was used to predict the genomic terminus type, but Saba was unclassified by this method (8). Genes were predicted from Glimmer v3.0 and MetaGeneAnnotator v1.0 outputs (9, 10). Potential tRNA genes were detected with ARAGORN v2.36 (11). Putative Rho-independent terminators were assessed with TransTermHP v2.09 (12). Protein functions were predicted primarily with InterProScan v5.33-72 and BLAST v2.2.31 searched with a 0.001 maximum expectation value cutoff against the NCBI nonredundant and UniProtKB Swiss-Prot/TrEMBL databases (13–15). Putative transmembrane domains were analyzed with TMHMM v2.0 (16). Whole-genomic DNA sequence similarity with top-related phage was calculated using the progressiveMauve v2.4.0 algorithm (17). The tools mentioned here (run at default parameters unless otherwise stated) are all accessible in the Galaxy instance hosted by the Center for Phage Technology

Citation Nguyen J, Harb L, Moreland R, Liu M, Gill JJ, Ramsey J. 2019. Complete genome sequence of *Proteus mirabilis* siphophage Saba. Microbiol Resour Announc 8:e01094-19. https://doi.org/10.1128/MRA.01094-19.

Editor John J. Dennehy, Queens College

Address correspondence to Jolene Ramsey, jolenerr@tamu.edu.

Received 6 September 2019
Accepted 13 September 2019
Published 10 October 2019

at https://cpt.tamu.edu/galaxy-pub (18), and annotation was performed in the Apollo instance (19).

The 60,056-bp Saba siphophage genome has 48.8% G+C content. With 75 coding sequences predicted, Saba has 93.5% coding density. Within the nonredundant database, Saba is most similar to *Proteus* phage pPM_01 (50.94% nucleotide identity and 55 shared proteins, GenBank accession number KP063118) and *Proteus* phage PM87 (50.83% nucleotide identity and 48 shared proteins, GenBank accession number MG030346), as well as enterobacterial phage Chi (29.45% nucleotide identity and 40 proteins, GenBank accession number JX094499). Given these similarities to Chi and Chi-like viruses, Saba may also be flagellotropic.

Data availability. The genome sequence and associated data for phage Saba were deposited under GenBank accession number MN062188, BioProject number PRJNA222858, SRA run number SRR8892147, and BioSample number SAMN11408687.

ACKNOWLEDGMENTS

This work was supported by funding from the National Science Foundation (award DBI-1565146). Additional support came from the Center for Phage Technology (CPT), an Initial University Multidisciplinary Research Initiative supported by Texas A&M University and Texas AgriLife, and from the Department of Biochemistry and Biophysics of Texas A&M University.

We are grateful for the advice and support of the CPT staff.

This announcement was prepared in partial fulfillment of the requirements for BICH464 Bacteriophage Genomics, an undergraduate course at Texas A&M University.

REFERENCES

1. Drzewiecka D. 2016. Significance and roles of Proteus spp. bacteria in natural environments. Microb Ecol 72:741–758. https://doi.org/10.1007/s00248-015-0720-6.
2. Armbruster CE, Mobley HLT, Pearson MM. 2018. Pathogenesis of Proteus mirabilis infection. EcoSal Plus. https://doi.org/10.1128/ecosalplus.ESP-0009-2017.
3. Nzakizwanayo J, Hanin A, Alves DR, McCutcheon B, Dedi C, Salvage J, Knox K, Stewart B, Metcalfe A, Clark J, Gilmore BF, Gahan CGM, Jenkins ATA, Jones BV. 2016. Bacteriophage can prevent encrustation and blockage of urinary catheters by Proteus mirabilis. Antimicrob Agents Chemother 60:1530–1536. https://doi.org/10.1128/AAC.02685-15.
4. Adams MH. 1956. Bacteriophages. Interscience Publishers, Inc., New York, NY.
5. 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.
6. 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.
7. 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.
8. 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.
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. 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.
13. 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.
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. The UniProt Consortium. 2019. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 47:D506–D515. https://doi.org/10.1093/nar/gky1049.
16. Krogh A, Larsson B, Heijne von 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. 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.
18. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Grüning 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.
19. Lee E, Helt GA, Reese JT, Munoz-Torres MC, Childers CP, Buels RM, Stein L, Holmes IH, Elsik CG, Lewis SE. 2013. Web Apollo: a Web-based genomic annotation editing platform. Genome Biol 14:R93. https://doi.org/10.1186/gb-2013-14-8-r93.