AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Proteus mirabilis* Phage Mydo

Benjamin T. Jones,[a] Lauren Lessor,[a] Chandler O'Leary,[a] Jason Gill,[a] Mei Liu[a]

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

ABSTRACT *Proteus mirabilis* is a pathogen that has been linked to nosocomial infections. Studies on phages infecting *P. mirabilis* may provide therapeutics for infections caused by antibiotic-resistant strains of this pathogen. Here, we announce the complete genome sequence of a *P. mirabilis* myophage, Mydo, which is distantly related to *Escherichia coli* phage rv5.

Proteus mirabilis is a Gram-negative enteric pathogen that is linked to a variety of hospital-acquired illnesses (1). It is intrinsically resistant to nitrofurantoin and tetracycline (1) and has been reported to have developed resistance to extended-spectrum cephalosporins and coresistance to other antibiotics due to the production of β-lactamases (2, 3). The study of phages infecting *P. mirabilis* may lead to alternative treatments for these antibiotic-resistant strains.

Phage Mydo was isolated from a wastewater sample collected from College Station, TX, in 2013 using a *Proteus mirabilis* isolate as the host. Host bacteria were cultured on nutrient broth or agar (Difco) at 37°C with aeration. Phages were cultured and propagated using the soft agar overlay method (4). It was identified as a myophage using negative-stain transmission electron microscopy performed at the Texas A&M University Microscopy and Imaging Center, as described previously (5). Phage genomic DNA was prepared using a modified Promega Wizard DNA cleanup kit protocol, as described previously (5). Pooled indexed DNA libraries were prepared using the Illumina TruSeq Nano low-throughput (LT) kit, and the sequence was obtained from the Illumina MiSeq platform using the MiSeq V2 500-cycle reagent kit, following the manufacturer's instructions, producing 1,112,580 paired-end reads for the index containing the phage Mydo genome. FastQC 0.11.5 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to quality control the reads. The reads were trimmed with FASTX-Toolkit 0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/download.html) before being assembled using SPAdes 3.5.0 (6). Contig completion was confirmed by PCR using primers (5′-GGTGTCTGGTACGTTGGTTC-3′ and 5′-TGTGTGTGACAACGTACCTG-3′) facing off the ends of the assembled contig and Sanger sequencing of the resulting product, with the contig sequence manually corrected to match the resulting Sanger sequencing read. Glimmer 3.0 (7) and MetaGeneAnnotator 1.0 (8) were used to predict protein-coding genes with manual verification, and tRNA genes were predicted with ARAGORN 2.36 (9). Rho-independent terminators were identified via TransTermHP v2.09 (http://transterm.cbcb.umd.edu/). Sequence similarity searches were performed by BLASTp 2.2.28 (10) with a maximum expectation cutoff of 0.001 against the NCBI nonredundant (nr), UniProt Swiss-Prot (11), and TrEMBL databases. InterProScan 5.15-54.0 (12), LipoP (13), and TMHMM v2.0 (14) were used to predict protein function. All analyses were conducted at default settings via the CPT Galaxy (15) and Web Apollo (16) interfaces (https://cpt.tamu.edu/galaxy-pub).

Phage Mydo was assembled at 81-fold coverage to a complete genome of 145,127 bp. Mydo has a G+C content of 45%, which is higher than that of its host (39%) (17). At both the nucleotide level and protein level determined by BLAST against the NCBI nr/nucleotide database (E value, <0.001), Mydo is closely related to *Klebsiella*

Citation Jones BT, Lessor L, O'Leary C, Gill J, Liu M. 2019. Complete genome sequence of *Proteus mirabilis* phage Mydo. Microbiol Resour Announc 8:e01312-19. https://doi.org/10.1128/MRA.01312-19.

Editor John J. Dennehy, Queens College

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

Received 17 October 2019
Accepted 25 October 2019
Published 21 November 2019

phages vB_KpnM_BIS47 (GenBank accession number KY652726) and vB_KpnM_KB57 (GenBank accession number KT934943). Mydo shares 86% and 83% DNA similarity and 230 and 223 proteins (out of 264 total predicted proteins in Mydo) with phage vB_KpnM_BIS47 and phage vB_KpnM_KB57, respectively. With 77 shared proteins (determined via BLASTp; E value, <0.001), phage Mydo is also distantly related to *Escherichia coli* phage rv5 (GenBank accession number NC_011041) (18), placing it within a cluster of large, virulent myophages that infect Gram-negative hosts.

Data availability. The genome sequence of phage Mydo was submitted to GenBank as accession number MK024806. The associated BioProject, SRA, and BioSample accession numbers are PRJNA222858, SRR8771451, and SAMN11234226, respectively.

ACKNOWLEDGMENTS

This work was supported by funding from the National Science Foundation (awards EF-0949351 and 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 at Texas A&M University.

We are grateful for the advice and support of the CPT staff and the Texas A&M University Microscopy and Imaging Center.

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

REFERENCES

1. O'Hara CM, Brenner FW, Miller JM. 2000. Classification, identification, and clinical significance of Proteus, Providencia, and Morganella. Clin Microbiol Rev 13:534–546. https://doi.org/10.1128/CMR.13.4.534.
2. Stürenburg E, Mack D. 2003. Extended-spectrum beta-lactamases: implications for the clinical microbiology laboratory, therapy, and infection control. J Infect 47:273–295. https://doi.org/10.1016/s0163-4453(03)00096-3.
3. Luzzaro F, Perilli M, Amicosante G, Lombardi G, Belloni R, Zollo A, Bianchi C, Toniolo A. 2001. Properties of multidrug-resistant, ESBL-producing Proteus mirabilis isolates and possible role of beta-lactam/beta-lactamase inhibitor combinations. Int J Antimicrob Agents 17:131–135. https://doi.org/10.1016/s0924-8579(00)00325-3.
4. Adams MK. 1959. Bacteriophages. Interscience Publishers, Inc., New York, NY.
5. Gill JJ, Berry JD, Russell WK, Lessor L, Escobar-Garcia DA, Hernandez D, Kane A, Keene J, Maddox M, Martin R, Mohan S, Thorn AM, Russell DH, Young R. 2012. The Caulobacter crescentus phage phiCbK: genomics of a canonical phage. BMC Genomics 13:542. https://doi.org/10.1186/1471-2164-13-542.
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. 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.
8. 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.
9. 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.
10. 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.
11. The UniProt Consortium. 2018. UniProt: the universal protein knowledgebase. Nucleic Acids Res 46:2699. https://doi.org/10.1093/nar/gky092.
12. 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.
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. 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.
15. Cock PJ, Gruning BA, Paszkiewicz K, Pritchard L. 2013. Galaxy tools and workflows for sequence analysis with applications in molecular plant pathology. PeerJ 1:e167. https://doi.org/10.7717/peerj.167.
16. 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.
17. Falkow S, Ryman IR, Washington O. 1962. Deoxyribonucleic acid base composition of Proteus and Providence organisms. J Bacteriol 83:1318–1321.
18. Kropinski AM, Waddell T, Meng J, Franklin K, Ackermann HW, Ahmed R, Mazzocco A, Yates J, III, Lingohr EJ, Johnson RP. 2013. The host-range, genomics and proteomics of Escherichia coli O157:H7 bacteriophage rV5. Virol J 10:76. https://doi.org/10.1186/1743-422X-10-76.