AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Salmonella enterica* Serovar Newport Myophage Melville

Kailun Zhang,[a] Yicheng Xie,[a] Chandler J. O'Leary,[a] Mei Liu,[a] Jason J. Gill[a]

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

ABSTRACT Multiple antimicrobial-resistant strains of *Salmonella enterica* serovar Newport have been recorded. Study on phages infecting *S.* Newport may provide new therapeutics or diagnostics for this pathogen. Here, we describe the complete genome sequence of the T4-like phage Melville that uses *S.* Newport as one of its hosts.

The CDC listed *Salmonella enterica* serovar Newport as one of the top three *Salmonella* serotypes associated with human infections (1) and foodborne outbreaks (2, 3) in the United States. However, several strains of *S.* Newport display resistance to multiple classes of antimicrobials, including expanded-spectrum cephalosporins (4–6). The study of *S.* Newport phages will provide insights into the control of *Salmonella* bacteria.

Myophage Melville was isolated from a mixed wastewater sample from Austin, TX in August 2016 using *S.* Newport as the host. Host bacteria were cultured on tryptic soy broth or agar (Difco) at 37°C with aeration. The phage was isolated and propagated by the soft agar overlay method (7). Phage genomic DNA was prepared using a modified Promega Wizard DNA cleanup kit protocol (8). 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 667,982 paired-end reads for the index containing the phage genome. Quality-controlled (FastQC; http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) trimmed (FASTX-Toolkit 0.11.6; http://hannonlab.cshl.edu/fastx_toolkit/) reads were assembled using SPAdes 3.5.0 (9) into a contig at 132.8-fold coverage. The genome sequence was completed by PCR using primers (5′-TCTTCATAGCATGGGCACATATC-3′ and 5′-GGCGGGTGGTTTGAAGTAA-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. Protein-coding genes were predicted by Glimmer 3.0 (10) and MetaGeneAnnotator 1.0 (11), with manual correction. tRNA genes were analyzed using ARAGORN 2.36 (12). Protein functions were predicted based on sequence homology by BLASTp 2.2.28 (13). Conserved domain searches were conducted in InterProScan 5.15-5.40 (14). All analyses were conducted at default settings via the Center for Phage Technology (CPT) Galaxy (15) and WebApollo (16) interfaces (https://cpt.tamu.edu/).

The Melville genome (159,323 bp) has a G+C content of 37%, a level lower than that of *Salmonella* spp. (~50%) (17). Genes encoding dCMP hydroxymethylase that produce hydroxymethyl cytosine (HMC) DNA were identified, suggesting that there is protection from nucleases that degrade host DNA. Genes for holin (class III with a predicted single transmembrane domain [TMD] in an N-in, C-out topology), endolysin (murein hydrolase), and spanins for lysis of the host are distributed in the Melville genome. There is one self-splicing group I intron in the gene encoding thymidylate synthetase. The presence of the inner membrane protein *imm* (immunity) gene in Melville indicates the ability to exclude superinfecting phage.

Citation Zhang K, Xie Y, O'Leary CJ, Liu M, Gill JJ. 2019. Complete genome sequence of *Salmonella enterica* serovar Newport myophage Melville. Microbiol Resour Announc 8:e00255-19. https://doi.org/10.1128/MRA.00255-19.

Editor Catherine Putonti, Loyola University Chicago

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

Received 4 March 2019
Accepted 7 April 2019
Published 25 April 2019

Melville is a T4-like phage and belongs to the genus *S16virus*. Melville encodes homologs of the protector from prophage-induced early lysis genes *rIIA* and *rIIB*, as is common among T4-like phages. It shares 93.9% and 90.5% whole-genome DNA sequence identity by BLASTn with the *Salmonella* phage STML-198 (GenBank accession numbers NC_027344) and *Salmonella* phage vB_SenMS16 (S16; GenBank accession number NC_020416), respectively. As is the case with phage S16, Melville contains a tandem gene duplication of the predicted capsid vertex protein (GenBank accession numbers ATN93139 and ATN93140). The long tail fiber distal subunit of Melville (GenBank accession number ATN93217) has 76% identity with that of phage S16, which recognizes the outer membrane protein OmpC and has an unusually broad host range within the genus *Salmonella* (18).

Data availability. The genome sequence of phage Melville was deposited under GenBank accession number MF957259. The associated BioProject, SRA, and BioSample accession numbers are PRJNA222858, SRR8788210, and SAMN11259695, respectively.

ACKNOWLEDGMENTS

This work was supported by funding from the National Science Foundation (awards EF-0949351 and DBI-1565146) and from the National Cattlemen's Beef Association and Texas Beef Cattle. 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.

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

REFERENCES

1. CDC. 2013. Salmonella surveillance: annual summary. Centers for Disease Control and Prevention, Atlanta, GA.
2. Dev Kumar G, Micallef SA. 2017. Susceptibility of Salmonella enterica isolates from tomato farm environments to fatty acids naturally found on tomato fruit. Foodborne Pathog Dis 14:293–301. https://doi.org/10.1089/fpd.2016.2239.
3. Irvine WN, Gillespie IA, Smyth FB, Rooney PJ, McClenaghan A, Devine MJ, Tohani VK, Outbreak CT. 2009. Investigation of an outbreak of Salmonella enterica serovar Newport infection. Epidemiol Infect 137:1449–1456. https://doi.org/10.1017/S0950268809002416.
4. CDC. 2010. Investigation update: multi-state outbreak of human Salmonella Newport infections linked to raw alfalfa sprouts. Centers for Disease Control and Prevention, Atlanta, GA.
5. Gupta A, Fontana J, Crowe C, Bolstorff B, Stout A, Van Duyne S, Hoekstra MP, Whichard JM, Barrett TJ, Angulo FJ, National Antimicrobial Resistance Monitoring System PulseNet Working Group. 2003. Emergence of multidrug-resistant Salmonella enterica serotype Newport infections resistant to expanded-spectrum cephalosporins in the United States. J Infect Dis 188:1707–1716. https://doi.org/10.1086/379668.
6. Gebreyes WA, Thakur S. 2005. Multidrug-resistant Salmonella enterica serovar Muenchen from pigs and humans and potential interserovar transfer of antimicrobial resistance. Antimicrob Agents Chemother 49:503–511. https://doi.org/10.1128/AAC.49.2.503-511.2005.
7. Adams MK. 1959. Bacteriophages. Interscience Publishers, Inc., New York, NY.
8. 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.
9. 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.
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. 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.
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. 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. Hoffmann M, Luo Y, Lafon PC, Timme R, Allard MW, McDermott PF, Brown EW, Zhao S. 2013. Genome sequences of Salmonella enterica serovar Heidelberg isolates isolated in the United States from a multistate outbreak of human Salmonella infections. Genome Announc 1:e00004-12. https://doi.org/10.1128/genomeA.00004-12.
18. Marti R, Zurfluh K, Hagens S, Pianezzi J, Klumpp J, Loessner MJ. 2013. Long tail fibres of the novel broad-host-range T-even bacteriophage S16 specifically recognize Salmonella OmpC. Mol Microbiol 87:818–834. https://doi.org/10.1111/mmi.12134.