AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Escherichia coli* Siphophage Sciku

Isla Hernandez,[a] Micah Castillo,[a] Russell Moreland,[a] Mei Liu,[a] Jolene Ramsey[a]

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

ABSTRACT *Escherichia coli* is a Gram-negative bacterium that is found in humans and animals as both a commensal organism and a pathogen. This report describes the isolation of Sciku, a siphophage infecting *E. coli* 4s, with 73 protein-coding genes. Genome comparisons suggest that Sciku is related to phages within *Guernseyvirinae*.

Escherichia coli is a Gram-negative commensal bacterium found in the intestinal microflora of certain animals, including humans. However, not all strains are harmless, and they can cause diseases in humans, other mammals, and birds with intestinal or extraintestinal pathologies (1). Some *E. coli* strains carry virulence factors involved in the colonization of the intestinal tract required to develop pathology. Phage therapy is considered a viable strategy for treating *E. coli* infection in place of antibiotics, and to this end, we isolated bacteriophage Sciku (2).

Phage Sciku was isolated from filtered (0.2 μm) wastewater treatment sludge samples collected in College Station, Texas, using an *E. coli* 4s strain as the bacterial host (3). Both the phage and its host were grown aerobically at 37°C in Luria broth (BD), and standard soft-agar overlay methods were used in the isolation (4). Phage Sciku's genomic DNA was purified with a Promega Wizard DNA cleanup system with the shotgun library preparation modifications described by Summer (5). The sequencing library was prepared with a TruSeq Nano low-throughput kit and sequenced by Illumina MiSeq with v2 500-cycle chemistry. The 565,076 total sequence reads from the index containing the phage genome were quality controlled with FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc) and assembled with SPAdes v3.5.0 at 698.9-fold contig coverage after trimming using the FastX toolkit v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/) (6). PCR (forward primer, 5′-GGCACAGAAACCGTGTAATCT-3′; reverse primer, 5′-TGGACTCTGCCGCAAATATC-3′) and Sanger sequencing were used to close the phage genome. The Galaxy and Web Apollo instances hosted by the Center for Phage Technology (https://cpt.tamu.edu/galaxy-pub/) contain all the tools used for annotation; these were run at default parameters. For gene calling, we used GLIMMER v3.0 and MetaGeneAnnotator v1.0, along with ARAGORN v2.36 for the detection of tRNAs (7–11). Rho-independent termination sites were annotated from TransTermHP v2.09 (12). Gene function was predicted using InterProScan v5.33-72 and BLAST v2.2.31 at default settings with a maximum expectation value of 0.001 versus the NCBI nonredundant and UniProtKB Swiss-Prot and TrEMBL databases (13–15). Transmembrane domains were predicted using TMHMM v2.0 (16). Structural similarities were identified using the HHsuite v3.0 tool HHpred (multiple sequence alignment generation with HHblits using the ummiclus30_2018_08 database and modeling with the PDB_mmCIF70 database) (17). Genome-wide DNA sequence similarity was calculated with progressiveMauve v2.4.0 (18). For phage morphology, samples were negatively stained with 2% (wt/vol) uranyl acetate and viewed by transmission electron microscopy at the Texas A&M Microscopy and Imaging Center (19).

Citation Hernandez I, Castillo M, Moreland R, Liu M, Ramsey J. 2019. Complete genome sequence of *Escherichia coli* siphophage Sciku. Microbiol Resour Announc 8:e01052-19. https://doi.org/10.1128/MRA.01052-19.

Editor Simon Roux, DOE Joint Genome Institute

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

Received 26 August 2019
Accepted 30 August 2019
Published 19 September 2019

Sciku is a siphophage with a 43,130-bp genome, 50.1% G+C content, and 93.8% coding density. Our analysis assigned Sciku 73 protein-coding genes, with 34 ascribed a function, but no tRNAs. PhageTerm predicted headful packaging for Sciku, and the genome was opened in front of the small terminase subunit (20). Sciku has the highest similarity to the *Escherichia* phage VB_EcoS-Golestan (GenBank accession number MG099933) of *Guernseyvirinae*, with 56 similar unique proteins and 66.37% nucleotide identity. As seen in other *Guernseyvirinae*, Sciku has a large self-splicing intein with a Hint domain (InterProScan IPR036844) within one of its helicases (NCBI accession number QEG06907) (21).

Data availability. The genome sequence and associated data for phage Sciku were deposited under GenBank accession number MK931439, BioProject accession number PRJNA222858, SRA accession number SRR8893626, and BioSample accession number SAMN11414580.

ACKNOWLEDGMENTS

This work was supported by funding from the National Science Foundation (award number 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 Texas A&M University Department of Biochemistry and Biophysics.

We thank A. Letarov for the kind gift of the *Escherichia coli* strain 4s. 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. Bélanger L, Garenaux A, Harel J, Boulianne M, Nadeau E, Dozois CM. 2011. Escherichia coli from animal reservoirs as a potential source of human extraintestinal pathogenic E. coli. FEMS Immunol Med Microbiol 62:1–10. https://doi.org/10.1111/j.1574-695X.2011.00797.x.
2. Brüssow H. 2005. Phage therapy: the Escherichia coli experience. Microbiology 151:2133–2140. https://doi.org/10.1099/mic.0.27849-0.
3. Golomidova A, Kulikov E, Isaeva A, Manykin A, Letarov A. 2007. The diversity of coliphages and coliforms in horse feces reveals a complex pattern of ecological interactions. Appl Environ Microbiol 73:5975–5981. https://doi.org/10.1128/AEM.01145-07.
4. Adams MH. 1956. 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, 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.
8. 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.
9. 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.
10. 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.
11. 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.
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. 2018. UniProt: the universal protein knowledgebase. Nucleic Acids Res 46:2699. https://doi.org/10.1093/nar/gky092.
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 S-Z, Rau D, Kübler J, Lozajic M, Gabler F, Söding 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. 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.
19. 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.
20. 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.
21. Anany H, Switt AIM, De Lappe N, Ackermann H-W, Reynolds DM, Kropinski AM, Wiedmann M, Griffiths MW, Tremblay D, Moineau S, Nash JHE, Turner D. 2015. A proposed new bacteriophage subfamily: "Jerseyvirinae." Arch Virol 160:1021–1033. https://doi.org/10.1007/s00705-015-2344-z.