AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Escherichia coli* Bacteriophage U136B

A. R. Burmeister,[a,b] Denish Piya,[c] Paul E. Turner[a,b,d]

[a]Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut, USA
[b]BEACON Center for the Study of Evolution in Action, East Lansing, Michigan, USA
[c]Department of Bioengineering, University of California, Berkeley, Berkeley, California, USA
[d]Program in Microbiology, Yale School of Medicine, New Haven, Connecticut, USA

ABSTRACT We report the genome sequence of bacteriophage U136B, which is reliant on the lipopolysaccharide and the antibiotic efflux protein TolC for infection of *Escherichia coli* and is a useful model for studying trade-offs and trade-ups that shape evolution. Phage U136B has a 49,233-bp genome with 87 predicted genes.

Phage U136B is a useful model for studying evolutionary trade-offs and trade-ups that shape the evolution of bacterial populations (1, 2).

Bacteriophages were cultured on *Escherichia coli* growing on LB at 37°C with aeration. Bacteriophage U136B DNA was extracted using a phage DNA isolation kit (Norgen Biotek). The sequencing library was prepared using transposome-based tagmentation chemistry with the Nextera XT DNA kit (Illumina) and sequenced at the Yale Center for Genome Analysis on a MiSeq system to ~1,800× coverage. Sequences were randomly rarified to a target of 100× coverage to improve phage genome assembly (3). Illumina adaptor sequences were removed with Cutadapt version 2.6 (4). Sequences were trimmed for quality using Sickle version 1.33 (5) and assembled with SPAdes version 3.13 (6). The resulting assembly had 108× coverage of 49,350 bp. One contig (comprising 470 bp at 0.95× coverage) was omitted from further analysis. The phage U136B genome has a circularly permuted genome, as determined by terminal repeats of 127 bp. After removal of these 127 bp, a final contig contained 49,223 bp. The assembly was validated using BWA version 7.17 (7), with 99.7% of the sequences mapping back to the assembly, above the 90% threshold typically considered valid (3). The complete genome sequence was reverse complemented and reopened to be syntenic with bacteriophage TLS (NCBI reference sequence NC_009540). All tools were run with default parameters unless otherwise specified.

Genome annotation was conducted using the Galaxy (8) and Web Apollo (9) platforms for phage genome annotation (10). Structural gene prediction was completed using GLIMMER version 3.0 (11), MetaGeneAnnotator version 1.0 (12), and SixPack (13). Structural predictions were manually and individually confirmed based on the assessment of ribosomal binding sites, translation start/stop sites, and gene overlaps. Gene functions were predicted using BLASTp (14, 15), and putative functions were then manually and individually assigned upon review of the BLASTp results. One gene putatively encoding a tape measure chaperone via a slipper sequence (16) was manually identified and annotated using the ExPASy Translate tool (17) and BLASTp at the NCBI (18). No tRNA genes were found using either tRNAscan-SE version 2.0 (19) or ARAGORN (20). Of 87 predicted open reading frames (ORFs), 43 putative functional proteins and 44 hypothetical proteins were annotated. No ORFs had predicted integrase functions, consistent with prior results showing that phage U136B is strictly lytic (2). The GC content was determined to be 43% using GeeCee (21, 22).

Citation Burmeister AR, Piya D, Turner PE. 2021. Complete genome sequence of *Escherichia coli* bacteriophage U136B. Microbiol Resour Announc 10:e00030-21. https://doi.org/10.1128/MRA.00030-21.

Editor John J. Dennehy, Queens College CUNY

Address correspondence to A. R. Burmeister, alita.burmeister@yale.edu.

Received 24 February 2021
Accepted 10 March 2021
Published 1 April 2021

A whole-genome BLASTn search (18) indicated that phage U136B is in the *Tlsvirus* phage group, which includes *E. coli* bacteriophage TLS (GenBank accession number AY308796.1) (93% identity) and *E. coli* bacteriophage LL5 (NCBI reference sequence NC_047985.1) (94% identity), both of which are also reliant on the *E. coli* TolC efflux protein (23, 24) and have siphophage morphology with flexible tails (23, 25). Host range analysis of phage U136B indicates a limited host range of some, but not all, *E. coli* hosts (2).

Data availability. The annotated phage U136B genome has been deposited in NCBI GenBank under accession number MW598258. Original sequence reads have been deposited in the Sequence Read Archive (SRA) under SRA accession number SRR13337692 and BioProject accession number PRJNA688914.

ACKNOWLEDGMENTS

This work was supported by NIH grant R21AI144345 from the National Institute of Allergy and Infectious Diseases. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

1. Burmeister AR, Turner PE. 2020. Trading-off and trading-up in the world of bacteria-phage evolution. Curr Biol 30:R1120–R1124. https://doi.org/10.1016/j.cub.2020.07.036.
2. Burmeister AR, Fortier A, Roush C, Lessing AJ, Bender RG, Barahman R, Grant R, Chan BK, Turner PE. 2020. Pleiotropy complicates a trade-off between phage resistance and antibiotic resistance. Proc Natl Acad Sci U S A 117:11207–11216. https://doi.org/10.1073/pnas.1919888117.
3. Philipson CW, Voegtly LJ, Lueder MR, Long KA, Rice GK, Frey KG, Biswas B, Cer RZ, Hamilton T, Bishop-Lilly KA. 2018. Characterizing phage genomes for therapeutic applications. Viruses 10:188. https://doi.org/10.3390/v10040188.
4. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–11. https://doi.org/10.14806/ej.17.1.200.
5. Joshi NA, Fass JN. 2011. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files, v1.33. https://github.com/najoshi/sickle.
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. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
8. 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.
9. 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.
10. 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.
11. 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.
12. 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.
13. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res 47:W636–W641. https://doi.org/10.1093/nar/gkz268.
14. Cock PJ, Chilton JM, Gruning B, Johnson JE, Soranzo N. 2015. NCBI BLAST+ integrated into Galaxy. GigaScience 4:39. https://doi.org/10.1186/s13742-015-0080-7.
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. Xu J, Hendrix RW, Duda RL. 2004. Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. Mol Cell 16:11–21. https://doi.org/10.1016/j.molcel.2004.09.006.
17. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E, Grosdidier A, Hernandez C, Ioannidis V, Kuznetsov D, Liechti R, Moretti S, Mostaguir K, Redaschi N, Rossier G, Xenarios I, Stockinger H. 2012. ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res 40:W597–W603. https://doi.org/10.1093/nar/gks400.
18. NCBI Resource Coordinators. 2018. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 46:D8–D13. https://doi.org/10.1093/nar/gkx1095.
19. Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res 44:W54–W57. https://doi.org/10.1093/nar/gkw413.
20. 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.
21. Blankenberg D, Taylor J, Schenck I, He J, Zhang Y, Ghent M, Veeraraghavan N, Albert I, Miller W, Makova KD, Hardison RC, Nekrutenko A. 2007. A framework for collaborative analysis of ENCODE data: making large-scale analyses biologist-friendly. Genome Res 17:960–964. https://doi.org/10.1101/gr.5578007.
22. Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 16:276–277. https://doi.org/10.1016/s0168-9525(00)02024-2.
23. German GJ, Misra R. 2001. The TolC protein of *Escherichia coli* serves as a cell-surface receptor for the newly characterized TLS bacteriophage. J Mol Biol 308:579–585. https://doi.org/10.1006/jmbi.2001.4578.
24. Piya D, Lessor L, Koehler B, Stonecipher A, Cahill J, Gill JJ. 2020. Genome-wide screens reveal *Escherichia coli* genes required for growth of T1-like phage LL5 and V5-like phage LL12. Sci Rep 10:8058. https://doi.org/10.1038/s41598-020-64981-7.
25. Piya D, Lessor L, Liu M, Gill JJ. 2019. Complete genome sequence of enterotoxigenic *Escherichia coli* siphophage LL5. Microbiol Resour Announc 8:e00674-19. https://doi.org/10.1128/MRA.00674-19.