



# Complete Genome Sequence of *Salmonella enterica* Serovar Typhimurium Siphophage Siskin

Chandler O'Leary,<sup>a</sup> Yicheng Xie,<sup>a</sup> Rohit Kongari,<sup>a</sup> Jason J. Gill,<sup>a</sup> Mei Liu<sup>a</sup>

<sup>a</sup>Center for Phage Technology, Texas A&M University, College Station, Texas, USA

**ABSTRACT** Bacteriophage Siskin is a member of the  $\chi$ -like siphovirus phage cluster that infects *Salmonella enterica* serovar Typhimurium strain LT2. Here, we report the complete 58,476-bp sequence of the Siskin genome, provide confirmation of its genomic termini, and describe a potentially new class of holins and endolysins found in the lysis cassette.

*Salmonella enterica* serovar Typhimurium is a foodborne pathogen that causes millions of infections annually (1, 2). The bacteriophage Siskin could be used therapeutically or during food processing to combat this pathogen.

Siskin was isolated in 2016 from wastewater collected in Austin, TX, using *S. Typhimurium* strain LT2. Host bacteria were cultured on tryptic soy broth or agar (Difco) at 37°C with aeration, and phage were propagated using the soft agar overlay method (3). Phage genomic DNA was prepared using a modified Promega Wizard DNA cleanup kit protocol (4). Pooled indexed DNA libraries were prepared using the Illumina TruSeq Nano LT kit, and a sequence was obtained using the MiSeq V2 500-cycle reagent kit, following manufacturer's instructions, producing 539,431 reads for the index containing the phage genome. Reads were quality controlled in FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), trimmed with FASTX-Toolkit 0.11.6 ([http://hannonlab.cshl.edu/fastx\\_toolkit/download.html](http://hannonlab.cshl.edu/fastx_toolkit/download.html)), and assembled by SPAdes 3.5.0 (5) into a raw contig at 265.4-fold coverage. Protein-coding genes were predicted using GLIMMER 3.0 (6) and MetaGeneAnnotator 1.0 (7). tRNA genes were predicted with ARAGORN 2.36 (8). Protein functions were predicted primarily by sequence homology using BLASTp 2.2.28 (9), and the conserved domain was searched using InterProScan 5.15-54.0 (10). Rho-independent termination sites were identified via TransTerm (<http://transterm.cbcb.umd.edu/>). Transmembrane domain and topology were predicted by TMHMM 2.0 (11, 12). All analyses were conducted at default settings through the CPT Galaxy (13) and WebApollo (14) platforms (<https://cpt.tamu.edu/>).

Out of 74 predicted total proteins of Siskin, 68 are similar to those of phage  $\chi$  (GenBank accession no. [KM458633](https://www.ncbi.nlm.nih.gov/nuclseq/KM458633)) by BLASTp, with an E value of <0.001 (15, 16), indicating that Siskin is a member of the  $\chi$ -like phage cluster (17). Phage Siskin has 92.48% nucleotide similarity to phage Utah (18) (GenBank accession no. [KY014601](https://www.ncbi.nlm.nih.gov/nuclseq/KY014601)), as determined by progressiveMauve 2.4.0 (19). Like phages  $\chi$  (15, 16) and Utah (18), Siskin was predicted to package its DNA by a site-specific *cos* mechanism with 5' extended overhangs. With primers (5'-GGGCGGCTGAGAAAGAATTA-3' and 5'-CAACCGGTAAAA CCGTAA-3') facing off the ends of the assembled raw contig, PCR using the ligated genomic DNA was conducted, and the resulting product was Sanger sequenced. The obtained sequence was compared to the direct Sanger resequencing of the phage termini (20), using primer 5'-GGGCGGCTGAGAAAGAATTA-3' off one end and 5'-CAA CCGGTAAAAACCGTAA-3' off the other end. The Sanger sequencing comparison results confirmed that the 5' end of the Siskin chromosome is at a position identical to the 5'-overhang sequence determined for phages  $\chi$  and Utah. Based on the presence of

**Citation** O'Leary C, Xie Y, Kongari R, Gill JJ, Liu M. 2019. Complete genome sequence of *Salmonella enterica* serovar Typhimurium siphophage Siskin. *Microbiol Resour Announc* 8:e00188-19. <https://doi.org/10.1128/MRA.00188-19>.

**Editor** John J. Dennehy, Queens College

**Copyright** © 2019 O'Leary et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Mei Liu, [meiliu@tamu.edu](mailto:meiliu@tamu.edu).

**Received** 1 March 2019

**Accepted** 1 April 2019

**Published** 2 May 2019

identical terminal sequences (GGTGCGCAGAGC), the Siskin chromosome is predicted to have 12-bp 5' overhangs with the sequence 5'-GGTGCGCAGAGC-3', and the phage genome was opened at this point. The complete genome of 58,476 bp has a coding density of 93.91% and a G+C content of 56.55%.

The putative holin of Siskin is predicted to have an N-terminal transmembrane domain with N-out, C-in topology, and the putative endolysin is predicted to have two C-terminal transmembrane domains with N-in, C-in topology. These are novel findings and, if experimentally confirmed, reveal a new class of holins (class IV) and a new class of phage endolysins.

**Data availability.** The genome sequence of phage Siskin was deposited under GenBank accession no. [MH631453](https://doi.org/10.6026/9781107011147). The associated BioProject, SRA, and BioSample accession numbers are [PRJNA222858](https://doi.org/10.6026/9781107011147), [SRR8788536](https://doi.org/10.6026/9781107011147), and [SAMN11260697](https://doi.org/10.6026/9781107011147), 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 Texas A&M University Department of Biochemistry and Biophysics.

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

- Neidhardt FC. 1987. *Escherichia coli* and *Salmonella Typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, DC.
- Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, Jones JL, Griffin PM. 2011. Foodborne illness acquired in the United States—major pathogens. *Emerg Infect Dis* 17:7–15. <https://doi.org/10.3201/eid1701.P11101>.
- Adams MK. 1959. Bacteriophages. Interscience Publishers, Inc., New York, NY.
- 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>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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>.
- 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>.
- 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>.
- 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>.
- 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>.
- 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>.
- 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>.
- Sonnhammer EL, von Heijne G, Krogh A. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* 6:175–182.
- 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>.
- Lee E, Helt GA, Reese JT, Munoz-Torres MC, Childers CP, Buels RM, Stein L, Holmes IH, Elisk 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>.
- Lee JH, Shin H, Choi Y, Ryu S. 2013. Complete genome sequence analysis of bacterial-flagellum-targeting bacteriophage chi. *Arch Virol* 158:2179–2183. <https://doi.org/10.1007/s00705-013-1700-0>.
- Hendrix RW, Ko CC, Jacobs-Sera D, Hatfull GF, Erhardt M, Hughes KT, Casjens SR. 2015. Genome sequence of *Salmonella* phage chi. *Genome Announc* 3:e01229-14. <https://doi.org/10.1128/genomeA.01229-14>.
- Grose JH, Casjens SR. 2014. Understanding the enormous diversity of bacteriophages: the tailed phages that infect the bacterial family Enterobacteriaceae. *Virology* 468–470:421–443. <https://doi.org/10.1016/j.virol.2014.08.024>.
- Leavitt JC, Heitkamp AJ, Bhattacharjee AS, Gilcrease EB, Casjens SR. 2017. Genome sequence of *Escherichia coli* tailed phage Utah. *Genome Announc* 5:e01494-16. <https://doi.org/10.1128/genomeA.01494-16>.
- 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>.
- Casjens SR, Gilcrease EB. 2009. Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions. *Methods Mol Biol* 502:91–111. [https://doi.org/10.1007/978-1-60327-565-1\\_7](https://doi.org/10.1007/978-1-60327-565-1_7).