



Complete Annotated Genome Sequence of the *Salmonella enterica* Serovar Typhimurium LT7 Strain STK003, Historically Used in Gene Transfer Studies

 Julie Zaworski,^a Anne Guichard,^{a*} Alexey Fomenkov,^a Richard D. Morgan,^a  Elisabeth A. Raleigh^a

^aResearch Department, New England Biolabs, Ipswich, Massachusetts, USA

ABSTRACT The genome of *Salmonella enterica* serovar Typhimurium LT7 comprises a chromosome and two plasmids. One plasmid is very close to pSLT of *Salmonella* Typhimurium LT2; the second harbors a shufflon region. Prophage content is distinct: LT7 lacks Fels-1, while Gifsy-1 and Fels-2 show island-like divergence and likely programmed inversion, respectively.

The isolates *Salmonella enterica* subsp. *enterica* serovar Typhimurium LT7 and LT2 were early model organisms for the study of gene transfer (1–3). *Salmonella* taxonomy has varied (4); these two lysotypes differ in prophage content (5). Isolates of LT2 were found to be more genetically stable than those of LT7 (6), which remains a model for its mutator properties (6, 7). Our LT7 isolate, STK003, was MST1656 from Stanley Maloy.

A Pacific Biosciences RS II instrument and the SMRT Analysis pipeline were used for sequencing, contig assembly, and modification analysis as described in reference 8. Briefly, cells were grown in LB and harvested, and DNA was then isolated with phenol-chloroform (9). A SMRTbell library was constructed from 5 μ g sheared (Covaris g-TUBE, 5 to 10 kb) DNA; sequencing used polymerase 6, chemistry 4. A total of 1,964 Mb of sequence in 131,952 polymerase reads with mean subread lengths of 3,046 bp and an N_{50} subread length of 4,016 bp was obtained ($\sim 329\times$ coverage). Assembly (RS_HGAP_Assembly.3; smrtanalysis_2.3.0.140936.p5.167094) specified a 5-Mb expected genome size, minimum subread length (MSL) of 1,000 bp, minimum read accuracy of 0.80, and minimum polymerase read length (MPRL) of 1,000 bp (other parameters were at default settings). Three contigs resulted with sizes of 4,835,584 bp, 115,234 bp, and 109,793 bp. The overall structure was rechecked with RS_HGAP3, resulting in an MSL of 3,000 bp and an MPRL of 9,000 bp.

Annotation, manipulation, and analysis employed Geneious_9.1.8 (Biomatters) with the options “Annotate_from_Database” (resulting in a 100% DNA match to the genome sequences found under the GenBank accession numbers [AE006468](#), [AE006471](#), and [NZ_CP012929](#)), “Sequence,” and “change residue numbering.” Circularization employed reference 10 or the procedure described in reference 11, File S1. Coordinates were set to agree with the reference origins. The final circular polished (Quiver) assemblies were a chromosome (4,817,454 bp) and two plasmids (93,946 bp and 88,171 bp). The GC contents are, respectively, 52.2%, 53.1%, and 50%.

The 4.8-Mb chromosome of STK003 shows 98.8% nucleotide identity to LT2 (MAFFT [12]). One plasmid is a close relative of pSLT of LT2; the second most closely resembles p12-4374_96 from *Salmonella enterica* subsp. *enterica* serovar Heidelberg (GenBank accession number [NZ_CP012929](#)). LT7 lacks prophage Fels-1, which is expected from the lysotype (13, 14). Prophages Fels-2 and Gifsy-1 of the

Citation Zaworski J, Guichard A, Fomenkov A, Morgan RD, Raleigh EA. 2021. Complete annotated genome sequence of the *Salmonella enterica* serovar Typhimurium LT7 strain STK003, historically used in gene transfer studies. *Microbiol Resour Announc* 10:e01217-20. <https://doi.org/10.1128/MRA.01217-20>.

Editor Vincent Bruno, University of Maryland School of Medicine

Copyright © 2021 Zaworski et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

Address correspondence to Elisabeth A. Raleigh, raleigh@neb.com.

* Present address: Anne Guichard, INRAE, Agrocampus Ouest, Université de Rennes, IGEPP, Le Rheu, France.

Received 20 November 2020

Accepted 6 February 2021

Published 11 March 2021

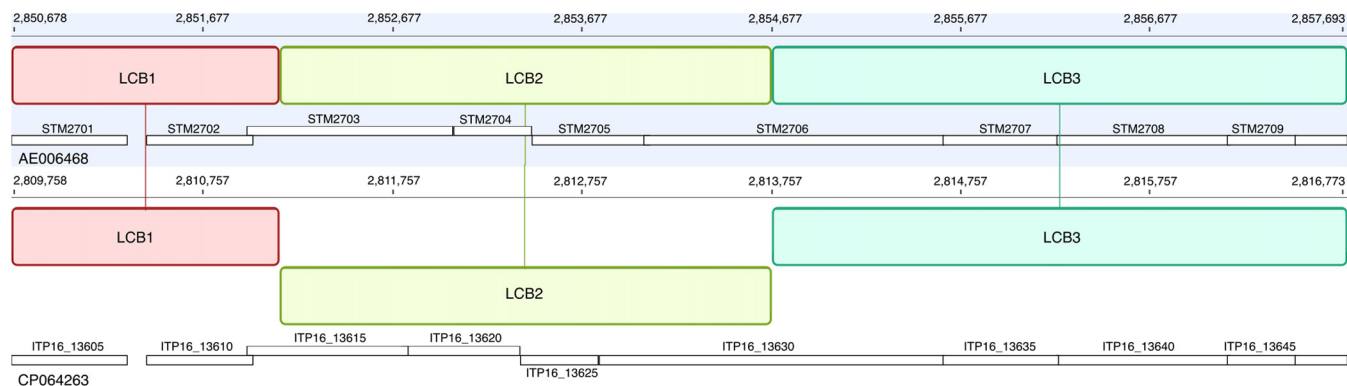


FIG 1 Inversion of candidate phage tail fiber coding sequence of Fels-2. Colored blocks represent segments of matching sequence (LCB), with genome coordinates above the blocks. Blocks above the horizontal lines have identity on the same strand; for the one below it, the bottom strand is identical to the top strand of its homolog. Accession numbers are below the horizontal lines; below that are open boxes labeled with locus tags of predicted coding DNA sequences (CDS). CDS STM2702 of LT2 (GenBank accession number [AE006468](#)) and its identical homolog, ITP16_13610 (*fin*) of LT7 ([CP064263](#)), adjoin the inversion. Alternative C-terminal segments are fused to identical N termini for STM2703 and ITP16_13615; similarly, identical N termini of STM2706 and ITP16_13630 are fused to alternative C termini. (We observe that notes in the LT2 sequence misleadingly describe STM2703 as similar to *pin* of e14; a more extensive description of STM2702 accurately notes similarity to site-specific recombinases and invertases.)

two strains differ in important respects. Gifsy-1 carries a substitution of ~7.5 kb between LT2 loci STM2620 and STM2630, likely from a different prophage (e.g., serovar Newport [GenBank accession number [CP025230](#)]). In contrast, Gifsy-2 is almost identical to the LT2 copy. Inversion of a Fels-2 tail fiber gene (Fig. 1; from the Geneious_Prime_2020.2.4 implementation of Mauve [15]) likely was mediated by *pin* homolog ITP16_13610 (LT2 locus tag STM2702) (16). Alternative host ranges are conferred by similar inversions mediated by *cin* of phage P1 and *gin* of phage Mu (17) and possibly *pin* of the e14 prophage-like element. The second plasmid harbors an ~2-kb region related to shufflons of IncI conjugal plasmids, which employ site-specific recombination to generate variation in pilus protein adhesion specificity (18, 19).

The methylation phenotype (from RS_Modification_and_Motif_Analysis) and genotype agree with activities expected for the serovar (Table 1), confirming functionality of methyltransferase (MTase) open reading frames (ORFs). No major single nucleotide polymorphisms (SNPs) were found in the restriction-modification (R-M) systems. Partial activity of M.SenLT2IV orthologs was previously reported for LT2 and other *Salmonella enterica* serovars (20).

Data availability. The genome sequences described here have been deposited under GenBank accession numbers [CP064263](#) to [CP064265](#) (assembly number [ASM1547561v1](#)) (PGAP annotated [21, 22]) and SRA numbers [SRR12788408](#) and [SRR12788409](#).

TABLE 1 Methylation specificity of DNA MTases in *Salmonella enterica* subsp. *enterica* serovar Typhimurium LT7^a

MTase ^b	Motif ^c	Type of modification	Sites with IPD > 2 (%) ^d
M.SenLT7I (StyLT)	5' CAG*AG 3'	m6A	99.2
M.SenLT7III (StySA)	5' GATC*AG 3'	m6A	99.3
M.SenLT7II (StySB)	5' G*AGNNNNNNRTAYG 3'	m6A	98.8
	3' CTCNNNNNNY*ATRC 5'		100
M.Dam	5' G*ATC 3'	m6A	99.7
M.Dcm	5' C*CWGG 3'	(m5C)	0.3
M.SenLT7IV	5' ATGC*AT 3'	m6A	54.9

^a The PacBio technology does not detect m5C accurately.

^b For MTases associated with restriction activities (R-M), synonyms found in the restriction literature are in parentheses.

^c Modified sequence motifs (* precedes the methylated base).

^d The percentage of recognition sites that were methylated. IPD, interpulse duration (23).

ACKNOWLEDGMENTS

We thank Stanley Maloy and Anca Segall (San Diego State University) for *Salmonella* strains and phages and advice and training in their use.

REFERENCES

- Miyake T. 1962. Exchange of genetic material between *Salmonella* Typhimurium and *Escherichia coli* K-12. *Genetics* 47:1043–1052. <https://doi.org/10.1093/genetics/47.8.1043>.
- Colson AM, Colson C, Van Pel A. 1969. Host-controlled restriction mutants of *Salmonella typhimurium*. *J Gen Microbiol* 58:57–64. <https://doi.org/10.1099/00221287-58-1-57>.
- Colson C, Colson AM. 1971. A new *Salmonella typhimurium* DNA host specificity. *J Gen Microbiol* 69:345–351. <https://doi.org/10.1099/00221287-69-3-345>.
- Brenner FW, Villar RG, Angulo FJ, Tauxe R, Swaminathan B. 2000. *Salmonella* nomenclature. *J Clin Microbiol* 38:2465–2467. <https://doi.org/10.1128/JCM.38.7.2465-2467.2000>.
- Zinder ND, Lederberg J. 1952. Genetic exchange in *Salmonella*. *J Bacteriol* 64:679–699. <https://doi.org/10.1128/JB.64.5.679-699.1952>.
- Gong J, Liu WQ, Liu GR, Chen F, Li JQ, Xu GM, Wang L, Johnston RN, Eisenstark A, Liu SL. 2007. Spontaneous conversion between *mutL* and 6 bp Δ *mutL* in *Salmonella typhimurium* LT7: association with genome diversification and possible roles in bacterial adaptation. *Genomics* 90:542–549. <https://doi.org/10.1016/j.ygeno.2007.06.009>.
- Koch WH, Henrikson E, Eisenstadt E, Cebula TA. 1995. *Salmonella typhimurium* LT7 and LT2 strains carrying the *imp* operon on Colla. *J Bacteriol* 177:1903–1905. <https://doi.org/10.1128/jb.177.7.1903-1905.1995>.
- Stenmark B, Harrison OB, Eriksson L, Anton BP, Fomenkov A, Roberts RJ, Tooming-Klunderud A, Bratcher HB, Bray JE, Thulin-Hedberg S, Maiden MCJ, Mölling P. 2020. Complete genome and methylome analysis of *Neisseria meningitidis* associated with increased serogroup Y disease. *Sci Rep* 10:3644. <https://doi.org/10.1038/s41598-020-59509-y>.
- Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Press, Cold Spring Harbor, NY.
- Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol* 16:2601. <https://doi.org/10.1186/s13059-015-0849-0>.
- Kingston AW, Roussel-Rossin C, Dupont C, Raleigh EA. 2015. Novel *recA*-independent horizontal gene transfer in *Escherichia coli* K-12. *PLoS One* 10:e0130813. <https://doi.org/10.1371/journal.pone.0130813>.
- Katoh K, Asimenos G, Toh H. 2009. Multiple alignment of DNA sequences with MAFFT. *Methods Mol Biol* 537:39–64. https://doi.org/10.1007/978-1-59745-251-9_3.
- McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S, Layman D, Leonard S, Nguyen C, Scott K, Holmes A, Grewal N, Mulvaney E, Ryan E, Sun H, Florea L, Miller W, Stoneking T, Nhan M, Waterston R, Wilson RK. 2001. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* 413:852–856. <https://doi.org/10.1038/35101614>.
- Liu GR, Edwards K, Eisenstark A, Fu YM, Liu WQ, Sanderson KE, Johnston RN, Liu SL. 2003. Genomic diversification among archival strains of *Salmonella enterica* serovar Typhimurium LT7. *J Bacteriol* 185:2131–2142. <https://doi.org/10.1128/jb.185.7.2131-2142.2003>.
- Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14:1394–1403. <https://doi.org/10.1101/gr.2289704>.
- Mehta P, Casjens S, Krishnaswamy S. 2004. Analysis of the lambdaoid prophage element ϵ 14 in the *E. coli* K-12 genome. *BMC Microbiol* 4:4. <https://doi.org/10.1186/1471-2180-4-4>.
- Grindley NDF, Whiteson KL, Rice PA. 2006. Mechanisms of site-specific recombination. *Annu Rev Biochem* 75:567–605. <https://doi.org/10.1146/annurev.biochem.73.011303.073908>.
- Gyohda A, Funayama N, Komano T. 1997. Analysis of DNA inversions in the shufflon of plasmid R64. *J Bacteriol* 179:1867–1871. <https://doi.org/10.1128/jb.179.6.1867-1871.1997>.
- Brouwer MSM, Tagg KA, Mevius DJ, Iredell JR, Bossers A, Smith HE, Partridge SR. 2015. Incl shufflons: assembly issues in the next-generation sequencing era. *Plasmid* 80:111–117. <https://doi.org/10.1016/j.plasmid.2015.04.009>.
- Pirone-Davies C, Hoffmann M, Roberts RJ, Muruvanda T, Timme RE, Strain E, Luo Y, Payne J, Luong K, Song Y, Tsai YC, Boitano M, Clark TA, Korch J, Evans PS, Allard MW. 2015. Genome-wide methylation patterns in *Salmonella enterica* Subsp. *enterica* Serovars. *PLoS One* 10:e0123639. <https://doi.org/10.1371/journal.pone.0123639>.
- Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res* 46:D851–D860. <https://doi.org/10.1093/nar/gkx1068>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, Korch J, Turner SW. 2010. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods* 7:461–465. <https://doi.org/10.1038/nmeth.1459>.