Complete Genome Sequence of *Salmonella enterica* Serovar Typhi Strain ISP2825

Gi Young Lee,ᵃ Jeongmin Songᵃ

ᵃDepartment of Microbiology and Immunology, Cornell University, Ithaca, New York, USA

ABSTRACT *Salmonella enterica* serovar Typhi ISP2825, isolated in 1983 from a Chilean patient, is one of the major *S.* Typhi strains used for research, along with strains Ty2, CT18, and H58. The complete genome sequence of ISP2825, consisting of a 4,774,014-bp circular chromosome, will help us understand typhoid pathogenesis and evolution.

Salmonella enterica serovar Typhi is the causative agent of the life-threatening systemic disease typhoid fever, which is a major cause of infection-mediated morbidity and mortality in countries of endemicity. Humans are the only known host of *S.* Typhi. Despite its narrow host specificity, *S.* Typhi has remained a highly successful pathogen since its emergence (1). Typhoid vaccines are available, but their efficacies are imperfect (2). Increasing antibiotic resistance of *S.* Typhi has been reported. Recently, multidrug-resistant (MDR) and extensively drug-resistant (XDR) strains have become the dominant *S.* Typhi variants (3).

There are four major *S.* Typhi strains extensively used for research that have helped advance our understanding of typhoid pathogenesis and evolution: Ty2 (4), CT18 (5), H58 (6), and ISP2825 (7–24). Complete genome sequences of these strains are available, except for ISP2825. Here, we report the complete genome sequence of *S.* Typhi ISP2825. In brief, *S.* Typhi ISP2825 was cultured in 2 ml Luria-Bertani broth (LB) overnight at 37°C, and its genomic DNA was extracted using the DNeasy blood and tissue kit (Qiagen, Germany), without processing additional fragmentation and size selection. The genomic DNA quality and quantity were monitored using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA) and a Qubit 4 fluorometer (Thermo Fisher Scientific). A genomic DNA library was prepared using a SQK-LSK110 ligation sequencing kit (Oxford Nanopore Technologies [ONT], UK), followed by sequencing with two MinION Flongle flow cells (R9.4.1) using MinKNOW v21.06.0 (ONT). The combined raw reads from two Flongle flow cells were used for base calling using Guppy v5.0.11 (ONT). Fastq files having Q scores of ≥8 were collected, filtered using NanoLyse v1.2.0 (25) to remove internal positive-control (DNA CS) sequences, and assessed for data quality using NanoPlot v1.38.1 (25) (Table 1). Next, the filtered reads were assembled using Flye v2.8.3 (26), with five iterations of the polishing step (–iterations 5), generating a single circular draft assembly. The resulting assembly was aligned using bwa v0.7.17 (https://github.com/lh3/bwa) and polished sequentially using Racon v1.4.22 (-m 8 -x -6 -g -8 -w 500) (27), Medaka v1.4.3 (https://github.com/nanoporetech/medaka) (-m r941_min_hac_g507), and Homopolish v0.2.3 (-s bacteria.msh -m R9.4.pkl) (28). The quality of the genome assembly was evaluated using QUAST v5.1.0 (29). Fasta_shift (https://github.com/b-brankovics/fasta_tools) was used to set the start position of the polished assembly according to that of *S.* Typhi CT18 (GenBank accession number GCF_000195995.1) and Ty2 (GCF_000007545.1). The complete genome sequence was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (30), and its predicted serotype

TABLE 1 Features of the complete genome sequence of *S.* Typhi ISP2825

Feature(s)	Value(s)
Filtered raw reads (Q score, ≥8)	
Total no. of reads	422,174
Total length (bp)	2,431,036,002
N_{50} (bp)	11,788
Complete genome	
Structure	Circular
Total length (bp)	4,774,014
G+C content (%)	52
Avg coverage (×)	511
Total no. of genes	4,688
No. of protein coding genes	4,310
No. of anticipated pseudogenes	267
No. of CRISPR arrays	1
RNA genes	
No. of rRNAs (5S, 16S, 23S)	8, 7, 7
No. of tRNAs	79
No. of noncoding RNAs	10
Computed antimicrobial resistance profile	None
Computed serotype	9:d:- (Typhi)

and antimicrobial resistance profiles were computed using SeqSero2 v1.2.1 (31) and AMRFinderPlus v3.10.5 (32), respectively (Table 1). Default options were used unless otherwise indicated. The complete genome sequence of ISP2825 consists of a 4,774,014-bp circular chromosome containing 4,688 genes, 4,310 protein coding genes, 267 anticipated pseudogenes, 22 rRNAs, and 79 tRNAs (Table 1). Overall, ISP2825 is similar to the other *S.* Typhi strains sequenced; notable differences are indicated in Fig. 1.

Data availability. The complete genome sequence of *S.* Typhi ISP2825 has been deposited at GenBank under accession number GCF_019645915.1 or CP080960.1, BioProject accession number PRJNA753482, BioSample accession number SAMN20695325, and Sequence Read Archive (SRA) accession number SRR15411315.
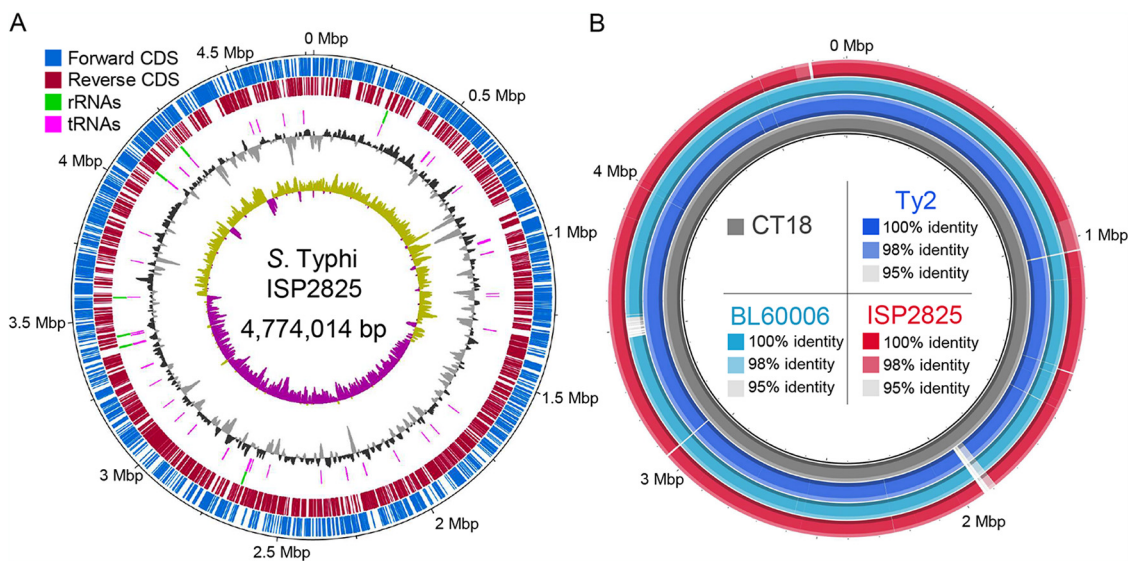

FIG 1 Circular maps summarizing features of *S.* Typhi ISP2825 (A) and comparative analysis of ISP2825, Ty2, BL60006, and CT18 (B). The ISP2825, Ty2, and BL60006 genomes were compared to the genome of CT18. Figure 1A and B were generated using DNAPlotter in Artemis (https://github.com/sanger-pathogens/Artemis) and BRIG v0.95 (https://sourceforge.net/projects/brig), respectively. CDS, the coding region of genes; BL60006, an XDR *S.* Typhi strain isolated from a Pakistani patient belonging to the H58 clade.

ACKNOWLEDGMENTS

REFERENCES

1. Dougan G, Baker S. 2014. Salmonella enterica serovar Typhi and the pathogenesis of typhoid fever. Annu Rev Microbiol 68:317–336. https://doi.org/10.1146/annurev-micro-091313-103739.

2. Syed KA, Saluja T, Cho H, Hsiao A, Shaikh H, Wartel TA, Mogasale V, Lynch J, Kim JH, Excler J-L, Sahastrabuddhe S. 2020. Review on the recent advances on typhoid vaccine development and challenges ahead. Clin Infect Dis 71:S141–S150. https://doi.org/10.1093/cid/ciaa504.

3. Klemm EJ, Shakoor S, Page AJ, Qamar FN, Judge K, Saeed DK, Wong VK, Dallman TJ, Nair S, Baker S, Shaheen G, Qureshi S, Yousafzai MT, Saleem MK, Hasan Z, Dougan G, Hasan R. 2018. Emergence of an extensively drug-resistant Salmonella enterica serovar Typhi clone harboring a promiscuous plasmid encoding resistance to fluoroquinolones and third-generation cephalosporins. mBio 9:e00105-18. https://doi.org/10.1128/mBio.00105-18.

4. Deng W, Liou S-R, Plunkett G, III, Mayhew GF, Rose DJ, Burland V, Kodoyianni V, Schwartz DC, Blattner FR. 2003. Comparative genomics of Salmonella enterica serovar Typhi strains Ty2 and CT18. J Bacteriol 185:2330–2337. https://doi.org/10.1128/JB.185.7.2330-2337.2003.

5. Parkhill J, Dougan G, James KD, Thomson NR, Pickard D, Wain J, Churcher C, Mungall KL, Bentley SD, Holden MTG, Sebaihia M, Baker S, Basham D, Brooks K, Chillingworth T, Connerton P, Cronin A, Davis P, Davies RM, Dowd L, White N, Farrar J, Feltwell T, Hamlin N, Haque A, Hien TT, Holroyd S, Jagels K, Krogh A, Larsen TS, Leather S, Moule S, O'Gaora P, Parry C, Quail M, Rutherford K, Simmonds M, Skelton J, Stevens K, Whitehead S, Barrell BG. 2001. Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18. Nature 413:848–852. https://doi.org/10.1038/35101607.

6. Wong VK, Baker S, Pickard DJ, Parkhill J, Page AJ, Feasey NA, Kingsley RA, Thomson NR, Keane JA, Weill F-X, Edwards DJ, Hawkey J, Harris SR, Mather AE, Cain AK, Hadfield J, Hart PJ, Thieu NTV, Klemm EJ, Glinos DA, Breiman RF, Watson CH, Kariuki S, Gordon MA, Heyderman RS, Okoro C, Jacobs J, Lunguya O, Edmunds WJ, Msefula C, Chabalgoity JA, Kama M, Jenkins K, Dutta S, Marks F, Campos J, Thompson C, Obaro S, MacLennan CA, Dolecek C, Keddy KH, Smith AM, Parry CM, Karkey A, Mulholland EK, Campbell JI, Dongol S, Basnyat B, Dufour M, Bandaranayake D, et al. 2015. Phylogeographical analysis of the dominant multidrug-resistant H58 clade of Salmonella Typhi identifies inter- and intracontinental transmission events. Nat Genet 47:632–639. https://doi.org/10.1038/ng.3281.

7. Galan JE, Curtiss R, III. 1991. Distribution of the invA, -B, -C, and -D genes of Salmonella typhimurium among other Salmonella serovars: invA mutants of Salmonella typhi are deficient for entry into mammalian cells. Infect Immun 59:2901–2908. https://doi.org/10.1128/iai.59.9.2901-2908.1991.

8. Haghjoo E, Galán JE. 2004. Salmonella Typhi encodes a functional cytolethal distending toxin that is delivered into host cells by a bacterial-internalization pathway. Proc Natl Acad Sci U S A 101:4614–4619. https://doi.org/10.1073/pnas.0400932101.

9. Haghjoo E, Galán JE. 2007. Identification of a transcriptional regulator that controls intracellular gene expression in Salmonella Typhi. Mol Microbiol 64:1549–1561. https://doi.org/10.1111/j.1365-2958.2007.05754.x.

10. Spanò S, Ugalde JE, Galán JE. 2008. Delivery of a Salmonella Typhi exotoxin from a host intracellular compartment. Cell Host Microbe 3:30–38. https://doi.org/10.1016/j.chom.2007.11.001.

11. Song J, Willinger T, Rongvaux A, Eynon EE, Stevens S, Manz MG, Flavell RA, Galan JE. 2010. A mouse model for the human pathogen Salmonella typhi. Cell Host Microbe 8:369–376. https://doi.org/10.1016/j.chom.2010.09.003.

12. Song J, Gao X, Galan JE. 2013. Structure and function of the Salmonella Typhi chimaeric A_2B_5 typhoid toxin. Nature 499:350–354. https://doi.org/10.1038/nature12377.

13. Chang S-J, Song J, Galan JE. 2016. Receptor-mediated sorting of typhoid toxin during its export from Salmonella Typhi-infected cells.

14. Cell Host Microbe 20:682–689. https://doi.org/10.1016/j.chom.2016.10.005.

14. Song J, Wilhelm CL, Wangdi T, Maira-Litran T, Lee S-J, Raetz M, Sturge CR, Mirpuri J, Pei J, Grishin NV, McSorley SJ, Gewirtz AT, Baumler AJ, Pier GB, Galan JE, Yarovinsky F. 2016. Absence of TLR11 in mice does not confer susceptibility to Salmonella Typhi. Cell 164:827–828. https://doi.org/10.1016/j.cell.2016.02.015.

15. Spanò S, Gao X, Hannemann S, Lara-Tejero M, Galán JE. 2016. A bacterial pathogen targets a host Rab-family GTPase defense pathway with a GAP. Cell Host Microbe 19:216–226. https://doi.org/10.1016/j.chom.2016.01.004.

16. Hannemann S, Galán JE. 2017. Salmonella enterica serovar-specific transcriptional reprogramming of infected cells. PLoS Pathog 13:e1006532. https://doi.org/10.1371/journal.ppat.1006532.

17. Fowler CC, Galán JE. 2018. Decoding a Salmonella Typhi regulatory network that controls typhoid toxin expression within human cells. Cell Host Microbe 23:65–76.e6. https://doi.org/10.1016/j.chom.2017.12.001.

18. Geiger T, Pazos M, Lara-Tejero M, Vollmer W, Galán JE. 2018. Peptidoglycan editing by a specific LD-transpeptidase controls the muramidase-dependent secretion of typhoid toxin. Nat Microbiol 3:1243–1254. https://doi.org/10.1038/s41564-018-0248-x.

19. Yang Y-A, Lee S, Zhao J, Thompson AJ, McBride R, Tsogtbaatar B, Paulson JC, Nussinov R, Deng L, Song J. 2018. In vivo tropism of Salmonella Typhi toxin to cells expressing a multiantennal glycan receptor. Nat Microbiol 3:155–163. https://doi.org/10.1038/s41564-017-0076-4.

20. Fowler CC, Stack G, Jiao X, Lara-Tejero M, Galán JE. 2019. Alternate subunit assembly diversifies the function of a bacterial toxin. Nat Commun 10:3684. https://doi.org/10.1038/s41467-019-11592-0.

21. Geiger T, Lara-Tejero M, Xiong Y, Galán JE. 2020. Mechanisms of substrate recognition by a typhoid toxin secretion-associated muramidase. Elife 9:e53473. https://doi.org/10.7554/eLife.53473.

22. Ahn C, Yang Y-A, Neupane DP, Nguyen T, Richards AF, Sim JH, Mantis NJ, Song J. 2021. Mechanisms of typhoid toxin neutralization by antibodies targeting glycan receptor binding and nuclease subunits. iScience 24:102454. https://doi.org/10.1016/j.isci.2021.102454.

23. Baldassarre M, Solano-Collado V, Balci A, Colamarino RA, Dambuza IM, Reid DM, Wilson HM, Brown GD, Mukhopadhyay S, Dougan G, Spanò S. 2021. The Rab32/BLOC-3-dependent pathway mediates host defense against different pathogens in human macrophages. Sci Adv 7:eabb1795. https://doi.org/10.1126/sciadv.abb1795.

24. Solano-Collado V, Colamarino RA, Calderwood DA, Baldassarre M, Spanò S. 2021. A small-scale shRNA screen in primary mouse macrophages identifies a role for the Rab GTPase Rab1b in controlling Salmonella Typhi growth. Front Cell Infect Microbiol 11:660689. https://doi.org/10.3389/fcimb.2021.660689.

25. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34:2666–2669. https://doi.org/10.1093/bioinformatics/bty149.

26. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8.

27. Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res 27:737–746. https://doi.org/10.1101/gr.214270.116.

28. Huang Y-T, Liu P-Y, Shih P-W. 2021. Homopolish: a method for the removal of systematic errors in Nanopore sequencing by homologous polishing. Genome Biol 22:95. https://doi.org/10.1186/s13059-021-02282-6.

29. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.

30. 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.

31. Zhang S, den Bakker HC, Li S, Chen J, Dinsmore BA, Lane C, Lauer AC, Fields
PI, Deng X. 2019. SeqSero2: rapid and improved Salmonella serotype deter-
mination using whole-genome sequencing data. Appl Environ Microbiol
85:e01746-19. https://doi.org/10.1128/AEM.01746-19.

32. Feldgarden M, Brover V, Gonzalez-Escalona N, Frye JG, Haendiges J,
Haft DH, Hoffmann M, Pettengill JB, Prasad AB, Tillman GE, Tyson GH,
Klimke W. 2021. AMRFinderPlus and the Reference Gene Catalog facilitate
examination of the genomic links among antimicrobial resistance, stress
response, and virulence. Sci Rep 11:12728. https://doi.org/10.1038/s41598-
021-91456-0.