Complete Genome Sequence of *Salmonella enterica* Serovar Typhimurium Strain SO3 (Sequence Type 302) Isolated from a Baby with Meningitis in Mexico

Pablo Vinuesa,[a] José L. Puente,[b] Edmundo Calva,[b] Mussaret B. Zaidi,[c] Claudia Silva[b]

Programa de Ingeniería Genómica, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico[a]; Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico[b]; Microbiology Research Laboratory, Hospital General O'Horan and Infectious Diseases Research Unit, Hospital Regional de Alta Especialidad de la Península de Yucatán, Mérida, Yucatán, Mexico[c]

The complete genome of *Salmonella enterica* serovar Typhimurium strain SO3 (sequence type 302), isolated from a fatal meningitis infection in Mexico, was determined using PacBio technology. The chromosome hosts six complete prophages and is predicted to harbor 51 genomic islands, including 13 pathogenicity islands (SPIs). It carries the *Salmonella* virulence plasmid (pSTV).

Address correspondence to Claudia Silva, csilvamex1@yahoo.com, or Pablo Vinuesa, vinuesa@ccg.unam.mx.

We report the complete genome sequence of *Salmonella enterica* subspecies *enterica* serovar Typhimurium strain SO3, which was isolated from the cerebrospinal fluid of a 1 month-old child who died of meningitis in Sonora, Mexico, in 2002. This strain was reported as part of an epidemiological surveillance program as SOHS 02-68 (1), and characterized by multilocus sequence typing as sequence type 302 (ST302) (2). This genotype was first described for human strains in Mexico (2), and later reported for African strains (http://mlst.warwick.ac.uk/mlst/dbs/Senterica). ST302 genotype is also closely related to ST313 human-invasive strains from Africa (3).

Genomic DNA was extracted by standard protocols (4) and sheared into ~10- to 20-kb fragments for PacBio library preparation and P6-C4 sequencing on one single-molecule real-time (SMRT) cell. The continuous-long-reads (CLRs) were assembled using the HGAP/Quiver protocol in SMRT Portal v2.3.0.140936.p4 (5), resulting in a two-contig assembly. These were circularized by trimming the terminal repeats with Minimus2 (6), and subjected to three consecutive rounds of CLRs remapping with the RS_Resequencing.1 module, yielding an assembly with 55× mean coverage. A final polishing step was performed by remapping quality-filtered Illumina HiSeq reads (PE 2 × 101 bp) onto the assembly using BWA (7), increasing its mean coverage to >450×. The alignment was passed to Pilon (8) to correct for indels and single nucleotide polymorphisms (SNPs). The size of the assembled genome is 5,006,803 bp, with a G+C content of 52 %, comprising a 4.9-Mb chromosome and a 93.8-kb *Salmonella* virulence plasmid (pSTV).

Gene calling and annotation was performed with a modified version of Prokka (9). A total of 4,980 genes, including 4,685 coding sequences (CDSs), and 11 pseudogenes were identified. Additionally, genes for 87 tRNAs, 22 rRNAs, and 1 tmRNA were annotated, plus 161 noncoding RNAs (ncRNAs), 3 clustered regularly interspaced short palindromic repeat (CRISPR) arrays, 5 riboswitches, and 442 signal peptides.

The annotation was manually curated, adding prophage predictions made by the PHAST server (10), and 51 genomic islands detected by IslandViewer3 (11). Among those, we identified and annotated 13 *Salmonella* pathogenicity islands (SPIs), including SPI-6, which encodes for a predicted type VI secretion system. Functional type III and type IV secretion systems were also predicted by querying the genome against effectiveDB (12). Six complete prophages and several phage remnants were located on the chromosome: ST104, Gifsy-2, ST64B, Gifsy-1, and the recently described P2-like phages ELPhiS (13) and FSL SP-004 (14). The phage repertoire and the pSTV sequence of strain SO3 were almost identical to the repertoire found in the genome of the other sequenced ST302 strain SO2 (15). A comparative genomics study of this and the other Mexican Typhimurium strains (16, 17) will be presented elsewhere.

Nucleotide sequence accession numbers. The complete sequences of the chromosome and the pSTV of *Salmonella* Typhimurium strain SO3 are available from GenBank under accession numbers CP014536 and CP014537, respectively.

REFERENCES

1. **Zaidi MB, Calva JJ, Estrada-Garcia MT, Leon V, Vazquez G, Figueroa G, Lopez E, Contreras J, Abbott J, Zhao S, McDermott P, Tollefson L.** 2008. Integrated food chain surveillance system for *Salmonella* spp. in Mexico. Emerg Infect Dis **14:**429–435. http://dx.doi.org/10.3201/eid1403.071057.

2. **Wiesner M, Zaidi MB, Calva E, Fernández-Mora M, Calva JJ, Silva C.** 2009. Association of virulence plasmid and antibiotic resistance determinants with chromosomal multilocus genotypes in Mexican *Salmonella enterica* serovar Typhimurium strains. BMC Microbiol **9:**131. http://dx.doi.org/10.1186/1471-2180-9-131.

3. **Kingsley RA, Msefula CL, Thomson NR, Kariuki S, Holt KE, Gordon MA, Harris D, Clarke L, Whitehead S, Sangal V, Marsh K, Achtman M, Molyneux ME, Cormican M, Parkhill J, MacLennan CA, Heyderman RS, Dougan G.** 2009. Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. Genome Res **19:**2279–2287. http://dx.doi.org/10.1101/gr.091017.109.

4. **Sambrook J, Russell DW.** 2001. Molecular cloning. A laboratory manual, **vol 1**, 3rd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

5. **Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J.** 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods **10:**563–569. http://dx.doi.org/10.1038/nmeth.2474.

6. **Treangen TJ, Sommer DD, Angly FE, Koren S, Pop M.** 2011. Next generation sequence assembly with AMOS. Curr Protoc Bioinformatics **33:**. http://dx.doi.org/10.1002/0471250953.bi1108s33.

7. **Li H, Durbin R.** 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics **26:**589–595. http://dx.doi.org/10.1093/bioinformatics/btp698.

8. **Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM.** 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One **9:**e112963. http://dx.doi.org/10.1371/journal.pone.0112963.

9. **Seemann T.** 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics **30:**2068–2069. http://dx.doi.org/10.1093/bioinformatics/btu153.

10. **Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS.** 2011. PHAST: a fast phage search tool. Nucleic Acids Res **39:**W347–W352. http://dx.doi.org/10.1093/nar/gkr485.

11. **Dhillon BK, Chiu TA, Laird MR, Langille MG, Brinkman FS.** 2013. IslandViewer update: improved genomic island discovery and visualization. Nucleic Acids Res **41:**W129–W132. http://dx.doi.org/10.1093/nar/gkt394.

12. **Eichinger V, Nussbaumer T, Platzer A, Jehl MA, Arnold R, Rattei T.** 2016. EffectiveDB—updates and novel features for a better annotation of bacterial secreted proteins and Type III, IV, VI secretion systems. Nucleic Acids Res **44:**D669–D674. http://dx.doi.org/10.1093/nar/gkv1269.

13. **Hanna LF, Matthews TD, Dinsdale EA, Hasty D, Edwards RA.** 2012. Characterization of the ELPhiS prophage from *Salmonella enterica* serovar Enteritidis strain LK5. Appl Environ Microbiol **78:**1785–1793. http://dx.doi.org/10.1128/AEM.07241-11.

14. **Moreno Switt AI, Orsi RH, den Bakker HC, Vongkamjan K, Altier C, Wiedmann M.** 2013. Genomic characterization provides new insight into *Salmonella* phage diversity. BMC Genomics **14:**481. http://dx.doi.org/10.1186/1471-2164-14-481.

15. **Silva C, Calva E, Puente JL, Zaidi MB, Vinuesa P.** 2016. Complete genome sequence of *Salmonella enterica* serovar Typhimurium strain SO2 (sequence type 302) isolated from an asymptomatic child in Mexico. Genome Announc **4(2):**e00253-16. http://dx.doi.org/10.1128/genomeA.00253-16.

16. **Calva E, Silva C, Zaidi MB, Sanchez-Flores A, Estrada K, Silva GG, Soto-Jiménez LM, Wiesner M, Fernández-Mora M, Edwards RA, Vinuesa P.** 2015. Complete genome sequencing of a multidrug-resistant and human-invasive *Salmonella enterica* serovar Typhimurium strain of the emerging sequence type 213 genotype. Genome Announc **3(3):**e00663-15. http://dx.doi.org/10.1128/genomeA.00663-15.

17. **Silva C, Calva E, Calva JJ, Wiesner M, Fernandez-Mora M, Puente JL, Vinuesa P.** 2015. Complete genome sequence of a human-invasive *Salmonella enterica* serovar Typhimurium strain of the emerging sequence type 213 harboring a multidrug resistance IncA/C plasmid and a *bla*CMY-2-Carrying IncF plasmid. Genome Announc **3(6):**e01323-15. http://dx.doi.org/10.1128/genomeA.01323-15.