Complete Genome Sequencing of a Multidrug-Resistant and Human-Invasive Salmonella enterica Serovar Typhimurium Strain of the Emerging Sequence Type 213 Genotype

Edmundo Calva,a Claudia Silva,a Mussaret B. Zaidi,b Alejandro Sanchez-Flores,c Karel Estrada,f Genivaldo G. Z. Silva,d Luz M. Soto-Jiménez,a Magdalena Wiesner,a* Marcos Fernández-Mora,a Robert A. Edwards,d,e,f Pablo Vinuesa,g

Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México; Microbiology Research Laboratory, Hospital General O’Horan and Infectious Diseases Research Unit, Hospital Regional de Alta Especialidad de la Península de Yucatán, Mérida, Yucatán, México; Unidad Universitaria de Apoyo Bioinformático, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México; Computational Science Research Center, San Diego State University, San Diego, California, USA; Department of Computer Science, San Diego State University, San Diego, California, USA; Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois, USA; Centro de Ciencias Genómicas, Programa de Ingeniería Genómica, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México.

* Present address: Magdalena Wiesner, Subdirección de Innovación, Dirección de Investigación en Salud Pública, Instituto Nacional de Salud, Bogotá, Colombia.

Salmonella enterica subsp. enterica serovar Typhimurium strain YU39 was isolated in 2005 in the state of Yucatán, Mexico, from a human systemic infection. The YU39 strain is representative of the multidrug-resistant emergent sequence type 213 (ST213) genotype. The YU39 complete genome is composed of a chromosome and seven plasmids.

Received 19 May 2015 Accepted 30 May 2015 Published 18 June 2015


Copyright © 2015 Calva et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported license.

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

An epidemiological surveillance program in Mexico showed that Salmonella enterica serovar Typhimurium is the most frequently isolated serovar from human infections (1). The multilocus sequence type 213 (ST213) was assigned to more than half of the Typhimurium population. This genotype was associated with food animal samples, lacked the Salmonella virulence plasmid, and carried multidrug resistance IncA/C plasmids (2, 3). Most of the systemic infections recorded during the surveillance period were caused by ST213 strains. Strain YU39 was isolated from the blood culture of an 8-year-old child displaying hepatomegaly and thrombocytopenia (3). This strain was studied for its capacity for conjugative transfer of the resistance blaBLM-2 gene through interactions between IncA/C and IncX1 plasmids (4).

Genomic DNA was extracted by standard protocols (5) and sheared into ~10-kb fragments for RSII-PacBio library preparation and P5-C3 sequencing. The continuous long read (CLR) data were assembled using the HGAP/Quiver protocol (SMRT portal version 2.2.0) (6). This resulted in an assembly containing eight contigs with ~70× genome coverage. A final polishing step was performed by remapping quality-filtered (7), 72-bp-long Illumina reads (n = 9,857,489, originating from a 200-bp paired-end library sequenced in a GAII instrument) and 454 GS FLX + single-end reads (n = 58,117; mean length, 418.04) onto the assembly using BWA (8), increasing its coverage to >150×. The aligned reads were converted to BAM format with SAMtools (9), and passed down to Pilon (10) to correct for small indels and SNPs. SMRTview analysis of CLRs mapped at the contig ends in the SMRT portal revealed their circular structure. Terminal repeats were trimmed with Minimus2 (11). The total size of the resulting genome assembly is 5,190,370 bp, with a G+C content of 51.94%, consisting of a 4.89-Mb chromosome and seven plasmids (156.3, 88.9, 42.2, 5.1, 4.8, 4.2, and 2.7 kb). Gene calling and annotation was performed with a modified version of the Flookka annotation pipeline (12). A total of 5,216 genes were identified, including 89 tRNAs, 22 rRNAs, 1 transfer-messenger RNA (tmRNA), 174 non-coding RNAs (ncRNAs), and 4,930 coding sequences (CDSs). Three CRISPR-CAS repeats and 453 signal peptides were annotated. The annotation was manually curated and enriched with predictions from the PHAST server (13) to annotate prophages, and IslandViewer to annotate genomic islands (14).

Five complete prophages were located on the chromosome: ST104, Gifsy-2, P88-like, ST64B, and Gifsy-2, as well as several phage remnants. The three large plasmids of 156.3, 88.9, and 42.2 kb were assigned to the multidrug resistance plnC/A/C, a phage-like plasmid, and the conjugative plnC/X1, respectively. This is the first complete genome sequence of a Mexican pathogenic and multidrug-resistant Salmonella Typhimurium strain.

Nucleotide sequence accession numbers. The complete genome sequences for the chromosome and the seven plasmids of Salmonella Typhimurium strain YU39 are available in GenBank under the accession numbers CP011428, CP011429, CP011430, CP011431, CP011432, CP011433, CP011434, and CP011435.

ACKNOWLEDGMENTS

This study was supported by grants from CONACyT (179946 to E.C. and 179133 to P.V.) and DGAPA/UNAM (IN201513 to E.C. and IN211814 to P.V.).

We are grateful to Francisco Javier Santana Estrada, Lucía Perezgasga
Ciscomani, and Freddy Campos for technical support; to Guilin Wang from the Yale Center for Genome Analysis, Yale University, for skillfully performing the PacBio sequencing; to Ricardo Grande from the Unidad Universitaria de Secuenciación Masiva y Bioinformática of the Instituto de Biotecnología, UNAM, for kindly performing the Illumina sequencing; and to Eugenio López, Santiago Becerra, Paul Gaytán, and Jorge Yáñez from the Unidad de Síntesis y Secuenciación of the Instituto de Biotecnología, UNAM.

REFERENCES


