AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Closed Genome Sequence of *Salmonella enterica* Serovar Richmond Strain CFSAN000191, Obtained with Nanopore Sequencing

Narjol González-Escalona,[a] Kuan Yao,[a] Maria Hoffmann[a]

[a]Division of Microbiology, Office of Regulatory Science, Center for Food Safety and Applied Nutrition, Food and Drug Administration, College Park, Maryland, USA

ABSTRACT Here we report the genome sequence of *Salmonella enterica* serovar Richmond strain CFSAN000191, isolated from tilapia from Thailand in 2005. The genome was determined by a combination of long-read and short-read sequencing. This strain was used for source tracking in a 2012 *Salmonella enterica* serovar Bareilly foodborne outbreak in the United States.

Salmonella enterica is one of the most important bacterial enteric pathogens and is implicated in foodborne illnesses worldwide (1). Among the many outbreaks of *Salmonella* infection occurring every year in the United States, there was an outbreak involving *Salmonella enterica* serovar Bareilly in 2012 that was linked to the consumption of a contaminated product containing raw yellowfin tuna (2). Single-nucleotide polymorphism (SNP) phylogenetic analysis using the whole-genome sequence (WGS) generated from 100 *S.* Bareilly strains (encompassing outbreak-related and non-outbreak-related strains) showed that the patients in the United States became infected with an *S.* Bareilly strain isolated from scraped tuna that was imported from a fishery in India for use in the production of spicy tuna sushi. The complete genome sequence of one representative strain (CFSAN000189) from that outbreak was obtained by PacBio sequencing (GenBank accession numbers CP006053 and CP006054) (2).

The genome of another *Salmonella* strain (CFSAN000191, reported as *S.* Bareilly) was sequenced to understand more about the diversity between *S.* Bareilly and *Salmonella enterica* serovar Richmond and to be used for future outbreak investigations. The strain was grown overnight in Luria-Bertani (LB) medium at 35°C, and the DNA was extracted with a DNeasy blood and tissue kit (Qiagen). The long reads for each strain were generated with MinION sequencing (Nanopore, Oxford, United Kingdom). The sequencing library was prepared with the RAD004 rapid sequencing kit. The sequencing library contained DNA fragmented randomly by a transposase present in the fragmentation mix of the RAD004 kit, rendering fragments >30 kb. This library was run in a FLO-MIN106 (R9.4.1) flow cell, according to the manufacturer's instructions, for 48 h. The sequencing output was 3 Gb (340,000 reads, but only reads above 5 kb were used for the downstream analyses, 132,158 reads) for an estimated average genome coverage of 360×. The short-read whole-genome sequence (WGS) for strain CFSAN000191 generated previously at a genome average coverage of 50× (2) was retrieved from the NCBI (SRA accession number SRR498369). The final genome sequence was achieved using the pipeline described previously (3). Briefly, the genome sequence was obtained with *de novo* assembly using Nanopore data and default settings in the Canu program v1.7 (4). A second assembly was generated using a SPAdes (5) hybrid assembly (with default settings) using both Nanopore and MiSeq data generated for the strain. The resulting assemblies from Canu were error corrected using Pilon (6) and the MiSeq data.

Received 25 October 2018 **Accepted** 2 November 2018 **Published** 13 December 2018

Citation González-Escalona N, Yao K, Hoffmann M. 2018. Closed genome sequence of *Salmonella enterica* serovar Richmond strain CFSAN000191, obtained with Nanopore sequencing. Microbiol Resour Announc 7:e01472-18. https://doi.org/10.1128/MRA.01472-18.

Editor Christina Cuomo, Broad Institute of MIT and Harvard University

Address correspondence to Narjol González-Escalona, narjol.gonzalez-escalona@fda.hhs.gov.

The final corrected assembly (FA) was generated by comparing the SPAdes hybrid and Canu-polished assemblies using Mauve (7). The two assemblies agreed in synteny and size, and therefore the SPAdes hybrid assembly was used as the FA. The FA consisted of a single contig of 4,726,630 bp (chromosome). The FA sequence was annotated using the NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAAP, https://www.ncbi.nlm.nih.gov/genome/annotation_prok). *In silico* multilocus sequence typing (MLST) analyses (https://enterobase.warwick.ac.uk/species/index/senterica) showed that CFSAN000191 belonged to sequence type 909 (ST909). *In silico* serotyping using SeqSero (8) (http://www.denglab.info/SeqSero), a tool to infer the serovar from the genes that determine antigenic structure, showed the strain to be *S.* Richmond (7:y:1,2) and not *S.* Bareilly (7:y:1,5), as initially suggested (2). The *S.* Bareilly and *S.* Richmond strains both belong to ST909. The GC content was 52.2%, similar to that of other *Salmonella* strains. Whole-genome SNP analysis, performed as described previously (9) but using the CFSAN000189 genome sequence as reference and 101 other *S.* Bareilly and *S.* Richmond genome sequences available at NCBI and used in a previous report (2), showed that the FA closed genome of this strain (CFSAN000191) was indistinguishable from the *de novo* MiSeq assembly for the same strain (JMMH00000000).

Data availability. The GenBank accession number for this genome sequence is CP032622, and the SRA accession number for the Nanopore run is SRR7941349.

ACKNOWLEDGMENTS

This study was supported by funding from MCMi challenge grant program proposal number 2018-646 and FDA Foods Program intramural funds.

REFERENCES

1. González-Escalona N, Hammack TS, Russell M, Jacobson AP, De Jesús AJ, Brown EW, Lampel KA. 2009. Detection of live *Salmonella* sp. cells in produce by a TaqMan-based quantitative reverse transcriptase real-time PCR targeting invA mRNA. Appl Environ Microbiol 75:3714–3720. https://doi.org/10.1128/AEM.02686-08.
2. Hoffmann M, Luo Y, Monday SR, González-Escalona N, Ottesen AR, Muruvanda T, Wang C, Kastanis G, Keys C, Janies D, Senturk IF, Catalyurek UV, Wang H, Hammack TS, Wolfgang WJ, Schoonmaker-Bopp D, Chu A, Myers R, Haendiges J, Evans PS, Meng J, Strain EA, Allard MW, Brown EW. 2016. Tracing origins of the *Salmonella* Bareilly strain causing a food-borne outbreak in the United States. J Infect Dis 213:502–508. https://doi.org/10.1093/infdis/jiv297.
3. Gonzalez-Escalona N, Haendiges J, Miller JD, Sharma SK. 2018. Closed genome sequences of two *Clostridium botulinum* strains obtained by Nanopore sequencing. Microbiol Resour Announc 7:e01075-18. https://doi.org/10.1128/MRA.01075-18.
4. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116.
5. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski 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.
6. 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. https://doi.org/10.1371/journal.pone.0112963.
7. 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.
8. Zhang S, Yin Y, Jones MB, Zhang Z, Deatherage Kaiser BL, Dinsmore BA, Fitzgerald C, Fields PI, Deng X. 2015. *Salmonella* serotype determination utilizing high-throughput genome sequencing data. J Clin Microbiol 53:1685–1692. https://doi.org/10.1128/JCM.00323-15.
9. Toro M, Retamal P, Ayers S, Barreto M, Allard M, Brown EW, Gonzalez-Escalona N. 2016. Whole-genome sequencing analysis of *Salmonella enterica* serovar Enteritidis isolates in Chile provides insights into possible transmission between gulls, poultry, and humans. Appl Environ Microbiol 82:6223–6232. https://doi.org/10.1128/AEM.01760-16.