Closed Genome Sequences of *Salmonella enterica* Serovar Javiana Isolates from Various Sources

Thomas Jurkiw,[a] Julie Haendiges,[a] Yan Luo,[b] Rebecca Bell,[a] Maria Hoffmann,[a] Jie Zheng[a]

aOffice of Regulatory Science, Center for Food Safety & Applied Nutrition, U.S. Food & Drug Administration, College Park, Maryland, USA
bOffice of Analytics and Outreach, Center for Food Safety & Applied Nutrition, U.S. Food & Drug Administration, College Park, Maryland, USA

ABSTRACT *Salmonella enterica* serovar Javiana is a major *Salmonella* serovar that causes human *Salmonella* infection in the United States. The complete genomic sequences of 9 *S.* Javiana isolates collected from food, environmental, and kratom sources in the United States were determined by hybrid assembly using Nanopore long-read sequencing and MiSeq short-read sequencing.

*S*almonella serovar Javiana is a predominant *Salmonella enterica* serovar in the United States, accounting for ~6% of all *Salmonella enterica* outbreaks in 2016 (1). Javiana infections have been traced back to a wide variety of food sources, such as tomato (2, 3), cheese (4), melons (5, 6), and herbs (7). Additionally, Javiana was one of the main serovars responsible for *Salmonella enterica* outbreaks in the medicinal herb kratom (8). Javiana is one of the few nontyphoidal *Salmonella* serovars harboring the genes encoding the Typhi cytolethal distending toxin (*pltA*, *pltB*, and *cdtB*), which causes G_2/M cell cycle arrest when injected into host cells (9, 10). Here, we report the complete genomic sequences of 9 *Salmonella enterica* Javiana isolates collected from food, environmental, and kratom sources in the United States. These genomes will aid in a better understanding of Javiana phylogeny.

All sequenced isolates were previously collected by the FDA except for CFSAN023356 and CFSAN023357, which were both obtained from the Minnesota Department of Health (Table 1). Isolates CFSAN086751 and CFSAN078221 were obtained from Kratom following FDA-BAM protocols (8). Isolates CFSAN038580, CFSAN055400, CFSAN038591, CFSAN048049, and CFSAN008899 were obtained using modified FDA-BAM protocols (11). These isolates were then grown overnight in tryptic soy broth (TSB) at 35°C and genomic DNA was extracted using the Maxwell RSC cultured cell DNA kit (Promega, Madison, WI) following the manufacturer's protocols. Short-read sequencing libraries were constructed from 100 ng of genomic DNA using the Illumina DNA prep (M) tagmentation kit and were sequenced on a MiSeq sequencer using MiSeq v3 kit with 2 × 250-bp paired-end chemistry (Illumina, San Diego, CA). The resulting reads were trimmed using Trimmomatic v0.36 and the SeqSero2 pipeline was used to confirm the serovar designation of the isolates (12, 13). The same DNA used for the MiSeq library prep was also used to construct libraries for Nanopore sequencing using the rapid barcoding sequencing kit (SQK-RBK004). The sequencing library was run in a FLO-MIN106 flow cell (R9.4.1) using a MinION Mk1C for either 48 or 72 h with live base calling using default settings (MinKNOW Core v4.3.12 and Guppy v5.0.16).

The complete genomic sequences for the 9 isolates were constructed using Unicycler to perform a hybrid assembly of the Nanopore and MiSeq data (14). Nanopore reads shorter than 5 kb were not used for constructing the hybrid assembly. The resulting assemblies were circularized and oriented to start at the *dnaA* gene. Separate assemblies were constructed from the Nanopore reads alone using Flye v2.9 (15) and compared to the Unicycler assemblies to ensure the two assemblies were in agreement. The hybrid

TABLE 1 Metadata and sequence assembly statistics for the 9 *Salmonella* serovar Javiana isolates reported in this study

Chromosome accession no.	CFSAN ID	Biosample ID	Collection source	Total no. of nanopore reads	N_{50} (bp)	Mean read length (bp)	Mean read quality	Read quality cutoff	Nanopore SRA accession no.	Total no. of Illumina reads	Illumina SRA accession no.	Genome size (bp)	GC content (%)	Hybrid assembly coverage (x)	Plasmid accession no.	Plasmid size (bp)
CP076097	CFSAN038580	SAMN04458018	Cucumber	102,307	9,821	5,018	11	7	SRR17335883	895,833	SRR17247936	4,583,372	53.42	181		
CP076098	CFSAN055400	SAMN05733259	Tomato	243,600	12,147	6,400	10	7	SRR17322490	623,820	SRR17247930	4,593,157	53.39	344		
CP076327	CFSAN048049	SAMN04933258	Water	200,697	13,998	7,185	10	7	SRR17335885	1,312,730	SRR17247935	4,583,367	53.43	402	CP076328	67,729
CP076329	CFSAN038591	SAMN04480334	Cucumber	240,000	13,249	7,106	10	7	SRR17335884	819,221	SRR17247933	4,583,372	53.42	526		
CP085052	CFSAN086751	SAMN10251118	Kratom	310,243	10,474	5,243	14	7	SRR17311396	530,394	SRR17247929	4,677,077	52.19	321		
CP085053	CFSAN078221	SAMN08737117	Kratom	222,881	10,562	5,236	14	7	SRR17322491	526,909	SRR18148071	4,611,726	52.20	248		
CP085054	CFSAN008899	SAMN02900330	Water	214,438	8,954	4,493	14	7	SRR17336095	230,424	SRR17247938	4,583,324	52.25	364	CP085055	67,729
CP086351	CFSAN023357	SAMN02989073	Turkey jerky	314,020	10,587	5,305	14	7	SRR17319836	743,859	SRR17247928	4,657,139	52.24	345	CP086352	6,194
CP086353	CFSAN023356	SAMN02989072	Turkey jerky	173,710	10,767	5,207	14	7	SRR17336057	626,712	SRR17247937	4,657,138	52.24	215		

assemblies were uploaded to NCBI and annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v5.3 (16). Three of the sequenced isolates had accompanying plasmid sequences: CFSAN048049, CFSAN008899, and CFSAN023357. The plasmids from CFSAN048049 and CFSAN008899 are identical ~68-kb plasmids containing *ccdAB* and *hicAB* type II toxin-antitoxin systems. The ~6.2 kb plasmid sequence from CFSAN023357 contains genes conferring tetracycline resistance [*tet*(A) and *tetR*].

Data availability. Accession numbers for all chromosome and plasmid sequences are listed in Table 1.

REFERENCES

1. Centers for Disease Control and Prevention. 2018. National enteric disease surveillance: Salmonella annual report, 2016. Centers for Disease Control and Prevention, Atlanta, GA.
2. Bennett SD, Littrell KW, Hill TA, Mahovic M, Behravesh CB. 2015. Multistate foodborne disease outbreaks associated with raw tomatoes, United States, 1990–2010: a recurring public health problem. Epidemiol Infect 143:1352–1359. https://doi.org/10.1017/S0950268814002167.
3. Hedberg CW, Angulo FJ, White KE, Langkop CW, Schell WL, Stobierski MG, Schuchat A, Besser JM, Dietrich S, Helsel L, Griffin PM, McFarland JW, Osterholm MT. 1999. Outbreaks of salmonellosis associated with eating uncooked tomatoes: implications for public health. The Investigation Team. Epidemiol Infect 122:385–393. https://doi.org/10.1017/S0950268899002393.
4. Hedberg CW, Korlath JA, D'Aoust JY, White KE, Schell WL, Miller MR, Cameron DN, MacDonald KL, Osterholm MT. 1992. A multistate outbreak of Salmonella Javiana and Salmonella Oranienburg infections due to consumption of contaminated cheese. JAMA 268:3203–3207. https://doi.org/10.1001/jama.1992.03490220047026.
5. Walsh KA, Bennett SD, Mahovic M, Gould LH. 2014. Outbreaks associated with cantaloupe, watermelon, and honeydew in the United States, 1973–2011. Foodborne Pathog Dis 11:945–952. https://doi.org/10.1089/fpd.2014.1812.
6. Centers for Disease Control. 2020. Outbreak of Salmonella infections linked to cut fruit. https://www.cdc.gov/salmonella/javiana-12-19.
7. Lehmacher A, Bockemuhl J, Aleksic S. 1995. Nationwide outbreak of human salmonellosis in Germany due to contaminated paprika and paprika-powdered potato chips. Epidemiol Infect 115:501–511. https://doi.org/10.1017/s0950268800058660.
8. Centers for Disease Control and Prevention. 2018. Multistate outbreak of Salmonella infections linked to kratom. Accessed 11/15/2021. Centers for Disease Control and Prevention, Atlanta, GA. https://www.cdc.gov/Salmonella/kratom-02-18/.
9. Mezal EH, Bae D, Khan AA. 2014. Detection and functionality of the CdtB, PltA, and PltB from Salmonella enterica serovar Javiana. Pathog Dis 72:95–103. https://doi.org/10.1111/2049-632X.12191.
10. Miller RA, Wiedmann M. 2016. The cytolethal distending toxin produced by nontyphoidal Salmonella serotypes Javiana, Montevideo, Oranienburg, and Mississippi induces DNA damage in a manner similar to that of Serotype Typhi. mBio 7. https://doi.org/10.1128/mBio.02109-16.
11. Bell RL, Zheng J, Burrows E, Allard S, Wang CY, Keys CE, Melka DC, Strain E, Luo Y, Allard MW, Rideout S, Brown EW. 2015. Ecological prevalence, genetic diversity, and epidemiological aspects of *Salmonella* isolated from tomato agricultural regions of the Virginia Eastern Shore. Front Microbiol 6:415. https://doi.org/10.3389/fmicb.2015.00415.
12. 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 determination using whole-genome sequencing data. Appl Environ Microbiol 85. https://doi.org/10.1128/AEM.01746-19.
13. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
14. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
15. 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.
16. 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.