AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Five Closed *Salmonella enterica* Genome Sequences from a 2017–2018 Multistrain, Multistate Kratom Outbreak

Hallie E. Rauch,[a] Julie Haendiges,[a] Maria Balkey,[a] Maria Hoffmann[a]

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

ABSTRACT We report here the closed genomes of *Salmonella enterica* strains from the 2017–2018 multistrain, multistate kratom outbreak using single-molecule real-time DNA sequencing. Four of the genomes consist of one circular chromosome, and the fifth has a circular chromosome and a single plasmid.

In January 2017, several cases of salmonellosis were reported to state and local health officials. Epidemiologic evidence suggested that the common source of these illnesses was kratom, a plant used as a homeopathic aid for chronic pain and opioid addiction. The Centers for Disease Control and Prevention (CDC) initially detected the outbreak as a cluster of *Salmonella enterica* serovar I4,[5],12:b:- through PulseNet, the CDC program for tracking outbreaks (https://www.cdc.gov/salmonella/kratom-02-18/index.html). From January 2017 to May 2018, a total of 199 cases were identified to be associated with this outbreak. These 199 identified cases resulted in 50 hospitalizations and spanned 41 states in the United States. The U.S. Food and Drug Administration (FDA) received *Salmonella* isolates obtained from kratom that were associated with the outbreak. As part of GenomeTrakr, the isolates were sequenced using Illumina technology, and subsequently, the serovars were predicted with SeqSero (1). In this study, we selected five of the *Salmonella* isolates from five different serovars and sequenced them using Pacific Biosciences (Menlo Park, CA) long-read sequence technology to establish high-quality reference genomes. One sample, CFSAN078398, was collected by the Utah Department of Health; the remaining four isolates were collected by the FDA.

The isolates were incubated overnight in tryptic soy broth (Becton, Dickinson, Franklin Lakes, NJ, USA), and genomic DNA was extracted with the Maxwell RSC cultured cell kit (Promega Corporation, Madison, WI). A 20-kb PacBio sample preparation protocol library was prepared, and size selection was performed with the Blue Pippin size selection system (Sage Science, Beverly, MA). The libraries were sequenced using P6-C4 chemistry on one or two single-molecule real-time (SMRT) cells with a 240-min collection time on the Pacific Biosciences RS II platform. The sequence coverage of the chromosomes ranged from 150 to 340× (Table 1). Analysis of the sequence reads was implemented using SMRT Analysis 2.3.0. *De novo* assembly of the reads was performed using the Hierarchical Genome Assembly Process 3 (HGAP3) program with default parameters (2). Overlapping regions identified at the end of the output assemblies (of chromosome and plasmids) were identified using Gepard 1.4 (3) and trimmed using an in-house script. The sequence data from each isolate resulted in one circular chromosome and one additional plasmid in genome CFSAN079094. The genomes were checked manually for even sequencing coverage. Afterward, the improved consensus sequence was uploaded in SMRT Analysis 2.3.0 to determine the final consensus and accuracy scores using the Quiver consensus algorithm. The sequencing statistics are listed in Table 1. The closed genome sequences were rotated to start at the *dnaA* gene. The assembled sequences were deposited at DDBJ/EMBL/GenBank and annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) 4.8 (4).

Citation Rauch HE, Haendiges J, Balkey M, Hoffmann M. 2020. Five closed *Salmonella enterica* genome sequences from a 2017–2018 multistrain, multistate kratom outbreak. Microbiol Resour Announc 9:e01395-19. https://doi.org/10.1128/MRA.01395-19.

Editor Catherine Putonti, Loyola University Chicago

Address correspondence to Maria Hoffmann, Maria.Hoffmann@fda.hhs.gov.

Received 6 November 2019
Accepted 22 December 2019
Published 16 January 2020

TABLE 1 Sequence identification and sequencing statistics

Nucleotide accession no.	CFSAN ID[a]	No. of SMRT cells	SRA accession no.	*In silico* serotyping result	GC content (%)	Coverage (×)	Mean read length (bp)	No. of reads	N_{50} read length (bp)
CP042443	CFSAN078398	2	SRX5107180 SRX5107169	Matopeni	52.1	315	13,497	145,341	23,263
CP042438	CFSAN079094	2	SRX5107185	Weltevreden	52.2	325	13,537	145,788	22,853
CP042439	pCFSAN079094		SRX5107168		48.8	500			
CP042442	CFSAN079101	1	SRX5107183	II 9,12:l,z28:5 or Javiana	52.2	150	14,239	71,570	24,247
CP042441	CFSAN079104	1	SRX5107157	Okatie or Newyork	52.3	166	14,839	73,773	26,601
CP042440	CFSAN079107	2	SRX5107154 SRX5107135	Corvallis or Chailey	52.1	340	13,417	152,564	24,102

[a] CFSAN ID, Center for Food Safety and Applied Nutrition identifier.

Each genome was scanned for antimicrobial resistance genes and chromosomal mutations using the Center for Genomic Epidemiology ResFinder 2.0 (5). All five genomes included the antimicrobial resistance gene *AAC(6′)-Iaa*, an aminoglycoside acetyltransferase gene (5). All genomes except CFSAN078398 contained a chromosomal mutation in the *parC* gene, a topoisomerase IV gene. This mutation is a missense transversion which resulted in a serine residue where a threonine residue would be in the original protein. The mutation is associated with fluoroquinolone resistance in *Salmonella* (6).

Isolate CFSAN079094 contained a single plasmid in addition to its circular chromosome. The plasmid was analyzed with CGE's PlasmidFinder 3.2 (7) and BLAST (8). The closest BLAST match, with 100% query coverage and 100% identity, for this plasmid was an uncharacterized plasmid associated with *Salmonella enterica* serovar Weltevreden. This plasmid is likely a member of the plasmid incompatibility group IncFII(S) and has an identity of 99.62%, representing one single-nucleotide polymorphism (SNP) (8).

Data availability. The complete genome sequences of the *Salmonella* isolates are publicly available in GenBank. The accession numbers are listed in Table 1.

ACKNOWLEDGMENT

This project was supported by the U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, Office of Regulatory Science.

REFERENCES

1. Zhang S, Yin Y, Jones MB, Zhang Z, Kaiser BLD, 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.
2. Chin C-S, 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. https://doi.org/10.1038/nmeth.2474.
3. Krumsiek J, Arnold R, Rattei T. 2007. Gepard: a rapid and sensitive tool for creating dotplots on genome scale. Bioinformatics 23:1026–1028. https://doi.org/10.1093/bioinformatics/btm039.
4. 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.
5. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother 67:2640–2644. https://doi.org/10.1093/jac/dks261.
6. Ling JM, Chan EW, Lam AW, Cheng AF. 2003. Mutations in topoisomerase genes of fluoroquinolone-resistant Salmonellae in Hong Kong. Antimicrob Agents Chemother 47:3567–3573. https://doi.org/10.1128/aac.47.11.3567-3573.2003.
7. Carattoli A, Zankari E, García-Fernández A, Larsen MV, Lund O, Villa L, Aarestrup FM, Hasman H. 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 58:3895–3903. https://doi.org/10.1128/AAC.02412-14.
8. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.