



Draft Genome Sequences of Shiga Toxin-Producing *Escherichia coli* O157:H7 Strains Recovered from a Major Production Region for Leafy Greens in California

 Beatriz Quiñones,^a Jaszemyn C. Yambao,^a Christopher J. Silva,^a Bertram G. Lee^a

^aU.S. Department of Agriculture, Agricultural Research Service, Western Regional Research Center, Produce Safety and Microbiology Research Unit, Albany, California, USA

ABSTRACT Shiga toxin-producing *Escherichia coli* O157:H7 is a foodborne pathogen and is responsible for outbreaks of human gastroenteritis. This report documents the draft genome sequences of nine O157:H7 cattle strains, which were identified to be PCR positive for a Shiga toxin gene but displayed different levels of functional toxin activity.

Shiga toxin-producing *Escherichia coli* (STEC) strains are enteric pathogens responsible for human gastroenteritis (1, 2). In some cases, human infections progress to hemolytic-uremic syndrome, a life-threatening disease resulting from the production of Shiga toxins (Stx). Serotype O157:H7 has been commonly associated with the development of severe disease symptoms. Cattle are a main reservoir for O157:H7 (3), and meat products are recognized as main sources of human infections (2). Recently, the consumption of leafy vegetables has been significantly linked to foodborne illness due to O157:H7 infections (4). Using a robust isolation method, STEC O157:H7 strains were recovered from cattle in a major agricultural region for leafy greens and were identified based on a typical STEC colony morphology on selective chromogenic medium and positive PCR tests for a *stx* gene (5). Further proteomic analyses indicated that all O157:H7 cattle strains were found to express the toxin receptor binding B-subunit (6, 7), but some were found not to have a functional toxin (7). The genome sequencing of nine O157:H7 cattle strains with various levels of toxin activity is reported here.

All STEC O157:H7 strains (Table 1) were streaked for isolation from a frozen stock culture on Luria-Bertani (LB) agar (Difco, Detroit, MI) at 37°C for 24 h. A single colony from each strain was further grown in LB broth (Difco) for 18 h with constant shaking (200 rpm) at 37°C. Genomic DNA of the O157:H7 strains was extracted with the Wizard genomic DNA purification kit (Promega Corp., Madison, WI) (8), sheared using a g-TUBE (Covaris, Inc., Woburn, MA), and quantified by fluorometric measurement using a Qubit 4.0 fluorometer (Invitrogen, Carlsbad, CA). Fifteen micrograms of the sample was used to prepare 20-kb SMRTbell libraries using the proprietary P6-C4 sequencing chemistry (Pacific Biosciences, Menlo Park, CA), according to the SMRTbell template prep kit 1.0 protocol (9). Single-molecule real-time (SMRT) sequencing was performed with the PacBio RS II platform (Pacific Biosciences) using the MagBead OneCellPerWell v1 collection protocol and 360-minute data collection mode (10). The sequencing reads were assembled using the PacBio Hierarchical Genome Assembly Process (HGAP; v3.0) and polished using Quiver in the SMRT Analysis v2.3.0 software, with default parameters (11), except for strain RM10645 and RM10646 assemblies (minimum polymerase read quality, 0.70; minimum seed read length, 5,000 bp). Methylation patterns were analyzed with RS_Modification_and_Motif_Analysis.1 using the SMRT Analysis software, with default settings. Annotations were performed using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP; version 4.8) (12). Genome sequence comparisons showed that all O157:H7 cattle strains were highly similar and had an average genome size

Citation Quiñones B, Yambao JC, Silva CJ, Lee BG. 2019. Draft genome sequences of Shiga toxin-producing *Escherichia coli* O157:H7 strains recovered from a major production region for leafy greens in California. *Microbiol Resour Announc* 8:e00644-19. <https://doi.org/10.1128/MRA.00644-19>.

Editor Frank J. Stewart, Georgia Institute of Technology

This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.

Address correspondence to Beatriz Quiñones, beatriz.quinones@ars.usda.gov.

Received 30 May 2019

Accepted 7 June 2019

Published 3 July 2019

TABLE 1 Genome characteristics and accession numbers of nine *Escherichia coli* O157:H7 strains from cattle from a leafy greens production region on the Central California Coast^a

Characteristic	Data for strain:									
	RM10024	RM10641	RM10645	RM10646	RM10649	RM10716	RM10718	RM10719	RM10720	
GenBank accession no.	QSMX000000000	QSMW000000000	QSMV000000000	QSMU000000000	QSMT000000000	QSMS000000000	QSMY000000000	QSMR000000000	QSMQ000000000	
SRA accession no.	SRX5679060	SRX5679061	SRX5679058	SRX5679059	SRX5679056	SRX5679057	SRX5679054	SRX5679055	SRX5679062	
Avg genome coverage (×) ^b	176	136	130	177	83	123	143	123	171	
N ₅₀ (bp) ^{b,c}	33,062	28,334	22,969	34,372	21,352	29,210	31,088	26,248	33,302	
Mean read length (bp) ^b	18,917	15,903	15,136	21,691	13,196	15,758	16,804	14,861	18,300	
Mapped read length of insert (bp) ^b	5,651	4,084	4,925	5,041	5,381	4,486	4,414	4,129	4,110	
No. of mapped reads ^b	66,346	60,966	73,290	81,452	42,756	52,256	58,706	57,840	64,823	
Genome size (bp) ^d	5,498,458	5,402,896	5,470,951	5,253,756	5,470,673	5,464,836	5,483,285	5,470,842	5,459,610	
G+C content (%) ^d	50.60	50.50	50.50	50.60	50.50	50.50	50.50	50.50	50.50	
Total no. of genes ^d	5,831	5,793	5,712	5,793	5,784	5,772	5,794	5,767	5,714	
No. of pseudogenes ^d	382	359	412	897	454	353	354	366	328	
stx gene subtype ^e	stx _{2c}	stx _{2c}	stx _{2c}	stx _{2c}	stx _{2c}	stx _{2c}	stx _{2c}	stx _{2c}	stx _{2c}	
Cytotoxicity assay result ^f	+	—	—	—	—	—	—	—	—	
stx screening test result ^g	+	+	+	+	+	+	+	+	+	

^a See reference 5.

^b Sequencing metrics were obtained using the PacBio RS II platform.

^c N₅₀, minimum contig length required to cover 50% of the total genome size.

^d Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline.

^e *In silico* serotype and stx subtype were confirmed by PCR (15, 16).

^f Cytotoxicity was determined with a fluorescent Vero cell-based assay (17, 18).

^g Sample enrichments were screened for stx genes by real-time PCR, and positive samples were further analyzed using selective chromogenic medium (5).

of 5,441,700 bp, with either 50.5% or 50.6% G+C content (Table 1). Analysis of methylation patterns revealed the DNA motifs **GATC**, **GAAABCC**, **CACNNNNNNNCTGG**, and **CCAGNNNNNNGTG** to be 98 to 99% methylated (an m6A modification occurred at the nucleotide in bold). The cattle strains harbored only one copy of the *stx_{2c}*-carrying prophage, which was inserted in the chromosomal *sbcB* gene, an insertion site commonly used by *stx_{2c}* prophages in virulent O157:H7 strains (13). Only strain RM10024 tested positive for cytotoxic activity (Table 1), but all other O157:H7 cattle strains tested negative due to the presence of the insertion sequence variant IS1203v (14) in the coding sequence of the catalytic A-subunit of Stx. The whole-genome sequencing information of O157:H7 cattle strains, recovered during a short time period and discrete sampling location in a major agricultural region, has provided an explanation for the variability in Stx activity and has set the foundation for future studies on the persistence of strains with an attenuated pathogenic potential.

Data availability. The assembled sequences and sequencing reads have been deposited at DDBJ/ENA/GenBank and the NCBI Sequence Read Archive, respectively, under the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

This material is based upon work supported in part by the U.S. Department of Agriculture (USDA), Agricultural Research Service, CRIS project number 2030-42000-051-00D and by the Trust Fund Cooperative Agreement number 58-5325-4-040 with Prosetta Bioconformatics, Inc., San Francisco, CA.

We thank Melissa Erickson-Beltran for critical comments on the manuscript.

REFERENCES

- Bolton DJ. 2011. Verocytotoxigenic (Shiga toxin-producing) *Escherichia coli*: virulence factors and pathogenicity in the farm to fork paradigm. *Foodborne Pathog Dis* 8:357–365. <https://doi.org/10.1089/fpd.2010.0699>.
- Karmali MA. 2018. Factors in the emergence of serious human infections associated with highly pathogenic strains of Shiga toxin-producing *Escherichia coli*. *Int J Med Microbiol* 308:1067–1072. <https://doi.org/10.1016/j.ijmm.2018.08.005>.
- Persad AK, Lejeune JT. 2014. Animal reservoirs of Shiga toxin-producing *Escherichia coli*. *Microbiol Spectr* 2:EHEC-0027-2014. <https://doi.org/10.1128/microbiolspec.EHEC-0027-2014>.
- Turner K, Moua CN, Hajmeer M, Barnes A, Needham M. 2019. Overview of leafy greens-related food safety incidents with a California link: 1996 to 2016. *J Food Prot* 82:405–414. <https://doi.org/10.4315/0362-028X.JFP-18-316>.
- Cooley MB, Jay-Russell M, Atwill ER, Carychao D, Nguyen K, Quiñones B, Patel R, Walker S, Swimley M, Pierre-Jerome E, Gordus AG, Mandrell RE. 2013. Development of a robust method for isolation of Shiga toxin-positive *Escherichia coli* (STEC) from fecal, plant, soil and water samples from a leafy greens production region in California. *PLoS One* 8:e65716. <https://doi.org/10.1371/journal.pone.0065716>.
- Fagerquist CK, Zaragoza WJ, Lee BG, Yambao JC, Quiñones B. 2019. Clinically-relevant Shiga toxin 2 subtypes from environmental Shiga toxin-producing *Escherichia coli* identified by top-down/middle-down proteomics and DNA sequencing. *Clin Mass Spectrom* 11:27–36. <https://doi.org/10.1016/j.clinms.2018.12.001>.
- Silva C, Lee B, Yambao J, Erickson-Beltran M, Quiñones B. 2019. Using nanospray liquid chromatography and mass spectrometry to quantitate Shiga toxin production in environmental *Escherichia coli* recovered from a major produce production region in California. *J Agric Food Chem* 67:1554–1562. <https://doi.org/10.1021/acs.jafc.8b05324>.
- Quiñones B, Yambao JC, Lee BG. 2017. Draft genome sequences of *Escherichia coli* O113: H21 strains recovered from a major produce production region in California. *Genome Announc* 5:e1203-17. <https://doi.org/10.1128/genomeA.01203-17>.
- Pacific Biosciences. 2018. Procedure & checklist—greater than 10 kb template preparation using AMPure PB beads. Pacific Biosciences, Menlo Park, CA. <https://www.pacb.com/wp-content/uploads/Procedure-Checklist-Greater-Than-10-kb-Template-Preparation-Using-AMPure-PB-Beads-1.pdf>.
- Pacific Biosciences. 2015. Procedure & checklist—preparing MagBeads for sequencing. Pacific Biosciences, Menlo Park, CA. <https://www.pacb.com/wp-content/uploads/2015/09/Procedure-Checklist-Preparing-MagBeads-for-Sequencing.pdf>.
- 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. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
- Tatusova T, Dicuccio M, Badretin 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>.
- Mellor GE, Sim EM, Barlow RS, D'Astek BA, Galli L, Chinen I, Rivas M, Gobius KS. 2012. Phylogenetically related Argentinean and Australian *Escherichia coli* O157 isolates are distinguished by virulence clades and alternative Shiga toxin 1 and 2 prophages. *Appl Environ Microbiol* 78:4724–4731. <https://doi.org/10.1128/AEM.00365-12>.
- Kusumoto M, Nishiya Y, Kawamura Y, Shinagawa K. 1999. Identification of an insertion sequence, IS1203 variant, in a Shiga toxin 2 gene of *Escherichia coli* O157:H7. *J Biosci Bioeng* 87:93–96. [https://doi.org/10.1016/S1389-1723\(99\)80014-0](https://doi.org/10.1016/S1389-1723(99)80014-0).
- Amézquita-López BA, Quiñones B, Lee BG, Chaidez C. 2014. Virulence profiling of Shiga toxin-producing *Escherichia coli* recovered from domestic farm animals in northwestern Mexico. *Front Cell Infect Microbiol* 4:7. <https://doi.org/10.3389/fcimb.2014.00007>.
- Quiñones B, Swimley MS, Narm KE, Patel RN, Cooley MB, Mandrell RE. 2012. O-antigen and virulence profiling of Shiga toxin-producing *Escherichia coli* by a rapid and cost-effective DNA microarray colorimetric method. *Front Cell Infect Microbiol* 2:61. <https://doi.org/10.3389/fcimb.2012.00061>.
- Quiñones B, Massey S, Friedman M, Swimley MS, Teter K. 2009. Novel cell-based method to detect Shiga toxin 2 from *Escherichia coli* O157:H7 and inhibitors of toxin activity. *Appl Environ Microbiol* 75:1410–1416. <https://doi.org/10.1128/AEM.02230-08>.
- Quiñones B, Swimley MS. 2011. Use of a Vero cell-based fluorescent assay to assess relative toxicities of Shiga toxin 2 subtypes from *Escherichia coli*. *Methods Mol Biol* 739:61–71. https://doi.org/10.1007/978-1-61779-102-4_6.