High-Quality Complete Genome Sequences of Three Bovine Shiga Toxin-Producing *Escherichia coli* O177:H- (*fliC*$_{H25}$) Isolates Harboring Virulent *stx2* and Multiple Plasmids

Haiqing Sheng,[a] Mingrui Duan,[a] Samuel S. Hunter,[b] Scott A. Minnich,[a] Matthew L. Settles,[c] Daniel D. New,[b] Jennifer R. Chase,[d] Matthew W. Fagnan,[b] Carolyn J. Hovde[a]

[a]Bi-State School of Food Science, University of Idaho, Moscow, Idaho, USA

[b]Institute for Bioinformatics and Evolutionary Studies (IBEST), University of Idaho, Moscow, Idaho, USA

[c]Bioinformatics Core, University of California, Davis, Davis, California, USA

[d]Department of Biology, Northwest Nazarene University, Nampa, Idaho, USA

ABSTRACT Shiga toxin-producing *Escherichia coli* (STEC) bacteria are zoonotic pathogens. We report here the high-quality complete genome sequences of three STEC O177:H- (*fliC*$_{H25}$) strains, SMN152SH1, SMN013SH2, and SMN197SH3. The assembled genomes consisted of one optical map-verified circular chromosome for each strain, plus two plasmids for SMN013SH2 and three plasmids for SMN152SH1 and SMN197SH3, respectively.

Shiga toxin (Stx)-producing *Escherichia coli* (STEC) can cause hemorrhagic colitis and hemolytic uremic syndrome (1–3). Cattle are major reservoirs of STEC and are the most common source for foodborne infections (4–6). During a longitudinal survey of STEC in healthy cattle, we isolated three non-O157 STEC strains that share phenotypic similarity with the O157 serogroup (7). Because these strains were nonmotile when isolated, DNA microarray for STEC serotyping (8) confirmed that these strains (SMN152SH1, SMN013SH2, and SMN197SH3, formerly named SH1, SH2 and SH3 respectively.) belonged to serogroup O177:H25. Based on these data we reassigned the serotype of these strains to O177:H- (*fliC*$_{H25}$: *fliC* genotype H25). Among the surveyed cattle in Washington and Idaho, US, and Alberta, Canada, the O177:H- serotype was more prevalent than the O157:H7 serotype (7). We report here the availability of three high-quality complete genomes of STEC O177:H- assembled using a hybrid approach that combines PacBio and Illumina reads.

Genomic DNA was prepared using a GenElute bacterial Genomic DNA kit (Sigma-Aldrich, St. Louis, MO) per the manufacturer's instructions. Whole-genome sequencing was carried out on a PacBio RS II (Pacific Biosciences, Menlo Park, CA) platform with 6-kb and 20-kb insert libraries by using P6 polymerase-binding and C4-sequencing kits for 240-min acquisition. Short-read sequencing was done at the IBEST Genomics Resources Core at the University of Idaho, using an Illumina MiSeq and v2 500-cyles kit (PE250) following construction of barcoded Illumina libraries using the IntegenX Apollo 324 PrepX ILM DNA library kit and custom Illumina TruSeq barcoded adapters. The Illumina reads were cleaned using HTStream (https://github.com/ibest/HTStream) to trim adapter sequences and low-quality ends and to overlap paired-end reads. Long reads were assembled with Canu v1.6 using default parameters (9). Gaps between contigs were sealed with GapBlaster (10). The small plasmids of SMN152SH1 and SMN197SH3 were assembled with Unicycler (11). The assembled genomes were oriented and trimmed using Circlator v1.5.2 (12), then corrected with Pilon v1.22 using cleaned Illumina reads (13). The resulting assemblies were verified using restriction enzyme NcoI whole-genome maps generated according to the OpGen protocol (OpGen, Inc., Gaithersburg, MD).

TABLE 1 Accession numbers and assembly metrics of the three annotated STEC O177:H-genomes

Strain	Plasmid	Accession no.	Genome size (bp)	GC content (%)	Plasmid size (bp)
SMN152SH1		CP024618	5,302,502	50.7	
	pO177A1	CP024617			88,896
	pO177B1	CP024616			48,983
	pO177C1	CP024615			6,675
SMN013SH2		CP023673	5,143,533	50.5	
	pO177A2	CP023674			87,524
	pO177B2	CP023675			46,664
SMN197SH3		CP024056	5,267,207	50.7	
	pO177A3	CP024055			88,839
	pO177B3	CP024054			48,982
	pO177C3	CP024053			6,675

The genomes were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (https://www.ncbi.nlm.nih.gov/genome/annotation_prok/). Features of the three O177:H- isolates include an stx_{2c} variant gene which is 100% identical to the stx_{2c} carried by a clinical isolate *E. coli* O177 (14); nearly identical pO157-like 88-kb and 48-kb plasmids, a 6.7-kb colicinogenic plasmid in SMN152SH1 and SMN197SH3; and chromosomally encoded AmpC type β-lactamase.

A detailed report on additional analysis of these genomes will be included in a future publication.

Accession number(s). Accession numbers and assembly metrics for each complete genome sequence are listed in Table 1.

REFERENCES

1. Edelman R, Karmali MA, Fleming PA. 1988. Summary of the International Symposium and Workshop on Infections Due to Verocytotoxin (Shiga-like toxin)-producing *Escherichia coli*. J Infect Dis 157:1102–1104. https://doi.org/10.1093/infdis/157.5.1102.
2. Karmali MA, Petric M, Lim C, Fleming PC, Arbus GS, Lior H. 1985. The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. J Infect Dis 151:775–782. https://doi.org/10.1093/infdis/151.5.775.
3. Smith HR, Scotland SM. 1988. Vero cytotoxin-producing strains of *Escherichia coli*. J Med Microbiol 26:77–85. https://doi.org/10.1099/00222615-26-2-77.
4. Blanco M, Blanco JE, Mora A, Dahbi G, Alonso MP, Gonzalez EA, Bernardez MI, Blanco J. 2004. Serotypes, virulence genes, and intimin types of Shiga toxin (verotoxin)-producing *Escherichia coli* isolates from cattle in Spain and identification of a new intimin variant gene (*eae-ξ*). J Clin Microbiol 42:645–651. https://doi.org/10.1128/JCM.42.2.645-651.2004.
5. Boerlin P, McEwen SA, Boerlin-Petzold F, Wilson JB, Johnson RP, Gyles CL. 1999. Associations between virulence factors of Shiga toxin-producing *Escherichia coli* and disease in humans. J Clin Microbiol 37:497–503.
6. Karmali MA. 1989. Infection by verocytotoxin-producing *Escherichia coli*. Clin Microbiol Rev 2:15–38. https://doi.org/10.1128/CMR.2.1.15.
7. Sheng H, Davis MA, Knecht HJ, Hancock DD, Van Donkersgoed J, Hovde CJ. 2005. Characterization of a Shiga toxin-, intimin-, and enterotoxin hemolysin-producing *Escherichia coli* ONT:H25 strain commonly isolated from healthy cattle. J Clin Microbiol 43:3213–3220. https://doi.org/10.1128/JCM.43.7.3213-3220.2005.
8. Lacher DW, Gangiredla J, Jackson SA, Elkins CA, Feng PC. 2014. Novel microarray design for molecular serotyping of Shiga toxin-producing *Escherichia coli* strains isolated from fresh produce. Appl Environ Microbiol 80:4677–4682. https://doi.org/10.1128/AEM.01049-14.
9. 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.
10. de Sá PHCG, Miranda F, Veras A, de Melo DM, Soares S, Pinheiro K, Guimarães L, Azevedo V, Silva A, Ramos RTJ. 2016. GapBlaster, a graphical gap filler for prokaryote genomes. PLoS One 11:e0155327. https://doi.org/10.1371/journal.pone.0155327.
11. 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.
12. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. Genome Biol 16:294. https://doi.org/10.1186/s13059-015-0849-0.
13. 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.
14. Gilmour MW, Tabor H, Wang G, Clark CG, Tracz DM, Olson AB, Mascarenhas M, Karmali MA, Mailman T, Ng LK. 2007. Isolation and genetic characterization of a coinfection of non-O157 Shiga toxin-producing *Escherichia coli*. J Clin Microbiol 45:3771–3773. https://doi.org/10.1128/JCM.01125-07.