Genome Sequences of 30 Escherichia coli O157:H7 Isolates Recovered from a Single Dairy Farm and Its Associated Off-Site Heifer-Raising Facility

Seon-Woo Kim, Jeffrey S. Karns, Jo Ann S. Van Kessel, Bradd J. Haley
Environmental Microbial and Food Safety Laboratory, Beltsville Area Research Center, Agricultural Research Services, United States Department of Agriculture, Beltsville, Maryland, USA

ABSTRACT
Cattle are the primary reservoir of Escherichia coli O157:H7, the most frequently isolated serotype of enterohemorrhagic E. coli infections among humans in North America. To evaluate the diversity of E. coli O157:H7 isolates within a single dairy herd, the genomes of 30 isolates collected over a 7-year period were sequenced.

Shiga toxin-producing Escherichia coli (STEC) is a major causative agent of foodborne gastroenteritis in North America, with serotype O157:H7 being responsible for approximately 36% of human clinical cases caused by STEC strains (1). The primary reservoirs of E. coli O157:H7 are known to be cattle and other ruminants, with carriage by other food animals, such as poultry and swine, rarely occurring (2). To evaluate the genomic diversity of E. coli O157:H7 within a single dairy farm and its associated off-site heifer-raising operation, we sequenced the genomes of 30 isolates recovered from these 2 farms and added the sequence data to the public database.

During the semiannual testing of an individual dairy farm and its associated heifer-raising operation for the presence of bacterial foodborne pathogens, at least 30 Shiga toxin-possessing E. coli O157:H7 isolates were recovered (Table 1). To prepare the isolates for genome sequencing, colonies were grown overnight in Luria-Bertani (LB) broth at 37°C, and one milliliter of this overnight growth was concentrated via centrifugation. DNA was extracted using the QIAcube sample preparation system (Qiagen, Hilden, Germany). Sequencing libraries were constructed using the Nextera XT library prep kit (Illumina, La Jolla, CA), which were sequenced using a high-output version 2.0 flow cell on a NextSeq 500 platform (Illumina). Raw reads were cleaned and trimmed using DeconSeq (3) and Trimmomatic (4) and assembled using SPAdes version 3.8.0 (5). After assembly, the genomes were analyzed in silico using the Center for Genomic Epidemiology webserver (http://www.genomicepidemiology.org/).

Based on the multilocus sequence typing (MLST) scheme developed by Wirth et al. (6), the 30 E. coli O157:H7 isolates consisted of 2 sequence types (STs), ST-11 (n = 26) and ST-5560 (n = 4). These STs have only a single allele difference between them. In a maximum parsimony phylogenetic analysis with 465 other E. coli O157:H7 genomes, the study isolates clustered into 3 divergent groups and one singleton (ARS-CC2204). The four ST-5560 isolates (ARS-CC2601, ARS-CC2248, ARS-CC2205, and ARS-CC2203) clustered closely together on a lineage nested within a larger clade composed mainly of ST-11 isolates, indicating that ST-5560 is a subclade of ST-11. Using the MLST scheme described by Jaureguy et al. (7), 29 isolates were identified as ST-628, but the single divergent isolate (ARS-CC2204) mentioned above was identified as ST-822. Isolates collected from the heifer farm clustered with those from the dairy farm, indicating that strains were likely transferred between these two locations.
Using the Harvest package (8), 2,151 single-nucleotide polymorphisms (SNPs) were detected among all strains excluding ARS-CC2204 (ST-628 strains only). When the ARS-CC2204 genome was added, the number of SNPs among the study isolates increased to 2,469 SNPs.

All genomes encoded sequences homologous to eae ( intimin), ehxA (plasmid-carried O157 enterohemolysin gene), and stx2 ( Shiga-toxin) virulence factors, all of which are integral in the enterohemorrhagic disease process of E. coli O157:H7. Transferable antibiotic resistance genes were not identified in any of the isolates.

Accession number(s). The genome sequences of these 30 E. coli isolates have been deposited in GenBank under the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

This project was supported by internal USDA, ARS research funding. The mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

REFERENCES

1. Pathela P, Klingler EJ, Guerry SL, Bernstein KT, Kerani RP, Liata L, Mark HD, Tabidez I, Rietmeijer CA, SsuN Working Group. 2015. Sexually transmitted infection clinics as safety net providers: exploring the role of categorial sexually transmitted infection clinics in an era of health care reform. Sex Transm Dis 42:286–293. https://doi.org/10.1097/OLQ .0000000000000255.

2. Ferens WA, Hovde CJ. 2011. Escherichia coli O157:H7: animal reservoir and sources of human infection. Foodborne Pathog Dis 8:465–487. https://doi.org/10.1089/fpd.2010.0673.

3. Schmieder R, Edwards R. 2011. Fast identification and removal of sequence contamination from genomic and metagenomic datasets. PLoS One 6:e17288. https://doi.org/10.1371/journal.pone.0017228.

4. 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.

5. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Strokin 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.

| Isolate | Serotype   | MLST as determined by methods of: | Isolation date (mo/day/yr) | Source | NCBI accession no. |
|---------|------------|-----------------------------------|---------------------------|--------|-------------------|
| ARS-CC2601 | O157:H7 | ST-5560 ST-628 | 3/14/2005 | Feces | NFRY00000000 |
| ARS-CC2203 | O157:H7 | ST-5560 ST-628 | 12/4/2005 | Fecal composite (postweaned calves) | NFRX00000000 |
| ARS-CC2204 | O157:H7 | ST-11 ST-822 | 12/4/2005 | Feces | NRFW00000000 |
| ARS-CC2205 | O157:H7 | ST-5560 ST-628 | 6/5/2006 | Fecal composite (postweaned calves) | NFRV00000000 |
| ARS-CC2207 | O157:H7 | ST-11 ST-628 | 12/4/2006 | Fecal composite (heifer grower) | NFRU00000000 |
| ARS-CC2248 | O157:H7 | ST-5560 ST-628 | 1/8/2008 | Feces | NRFW00000000 |
| ARS-CC5318 | O157:H7 | ST-11 ST-628 | 1/26/2010 | Feces (entering heifer) | NFRS00000000 |
| ARS-CC5541 | O157:H7 | ST-11 ST-628 | 12/4/2010 | Feces (entering heifer) | NRFQ00000000 |
| ARS-CC5557 | O157:H7 | ST-11 ST-628 | 12/4/2010 | Feces | NRFP00000000 |
| ARS-CC5525 | O157:H7 | ST-11 ST-628 | 6/3/2010 | Feces (calf at heifer grower) | NFRQ00000000 |
| ARS-CC5526 | O157:H7 | ST-11 ST-628 | 6/14/2010 | Feces (calf at heifer grower) | NFRN00000000 |
| ARS-CC5530 | O157:H7 | ST-11 ST-628 | 6/14/2010 | Feces | NFRM00000000 |
| ARS-CC5534 | O157:H7 | ST-11 ST-628 | 6/14/2010 | Feces | NFRN00000000 |
| ARS-CC5536 | O157:H7 | ST-11 ST-628 | 6/14/2010 | Fecal composite | NFRK00000000 |
| ARS-CC5539 | O157:H7 | ST-11 ST-628 | 6/14/2010 | Fecal composite | NFRJ00000000 |
| ARS-CC5546 | O157:H7 | ST-11 ST-628 | 6/14/2010 | Feces | NRFH00000000 |
| ARS-CC5548 | O157:H7 | ST-11 ST-628 | 6/14/2010 | Feces | NFRG00000000 |
| ARS-CC5549 | O157:H7 | ST-11 ST-628 | 6/14/2010 | Feces | NRFP00000000 |
| ARS-CC5550 | O157:H7 | ST-11 ST-628 | 6/14/2010 | Feces | NRFQ00000000 |
| ARS-CC5552 | O157:H7 | ST-11 ST-628 | 6/14/2010 | Feces | NRFD00000000 |
| ARS-CC5587 | O157:H7 | ST-11 ST-628 | 6/14/2010 | Feces | NRFC00000000 |
| ARS-CC5555 | O157:H7 | ST-11 ST-628 | 6/22/2010 | Feces | NFRB00000000 |
| ARS-CC5560 | O157:H7 | ST-11 ST-628 | 6/30/2010 | Feces (calf at heifer grower) | NFRK00000000 |
| ARS-CC6195 | O157:H7 | ST-11 ST-628 | 12/6/2010 | Fecal composite | NFGZ00000000 |
| ARS-CC6335 | O157:H7 | ST-11 ST-628 | 6/13/2011 | Fecal composite | NFGY00000000 |
| ARS-CC6338 | O157:H7 | ST-11 ST-628 | 6/13/2011 | Feces (calf at heifer grower) | NFGX00000000 |
| ARS-CC7094 | O157:H7 | ST-11 ST-628 | 9/24/2012 | Fecal composite | NFRW00000000 |
| ARS-CC7983 | O157:H7 | ST-11 ST-628 | 12/4/2012 | Fecal composite | NFRV00000000 |

Kim et al.

Volume 5 Issue 35 e00814-17

genomea.asm.org 2
6. Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MC, Ochman H, Achtman M. 2006. Sex and virulence in *Escherichia coli*: an evolutionary perspective. Mol Microbiol 60:1136–1151. https://doi.org/10.1111/j.1365-2958.2006.05172.x.

7. Jaureguy F, Landraud L, Passet V, Diancourt L, Frapy E, Guigon G, Carbonnelle E, Lortholary O, Clermont O, Denamur E, Picard B, Nassif X, Brisse S. 2008. Phylogenetic and genomic diversity of human bacteremic *Escherichia coli* strains. BMC Genomics 9:560. https://doi.org/10.1186/1471-2164-9-560.

8. Treangen TJ, Ondov BD, Koren S, Phillippy AM. 2014. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol 15:524. https://doi.org/10.1186/ PREACCEPT-2573980311437212.