




# Complete Genome Sequence of an Enterohemorrhagic *Escherichia coli* O111:H8 Strain Recovered from a Large Outbreak in Japan Associated with Consumption of Raw Beef

Tsuyoshi Sekizuka,<sup>a</sup>  Kenichi Lee,<sup>b</sup> Keiko Kimata,<sup>c</sup> Junko Isobe,<sup>c</sup> Makoto Kuroda,<sup>a</sup> Sunao Iyoda,<sup>b</sup> Makoto Ohnishi,<sup>b</sup> Tetsutaro Sata,<sup>c</sup> Masanori Watahiki<sup>c</sup>

<sup>a</sup>Pathogen Genomics Center, National Institute of Infectious Diseases, Tokyo, Japan

<sup>b</sup>Department of Bacteriology I, National Institute of Infectious Diseases, Tokyo, Japan

<sup>c</sup>Department of Bacteriology, Toyama Institute of Health, Toyama, Japan

**ABSTRACT** We present the complete genome sequence of an enterohemorrhagic *Escherichia coli* O111:H8 strain. This strain was isolated from a hemolytic-uremic syndrome patient and was responsible for a large outbreak associated with the consumption of raw beef in 2011.

Enterohemorrhagic *Escherichia coli* (EHEC) is a serious etiological agent of outbreaks of foodborne illness worldwide. More than 3,000 cases of EHEC infection are reported annually in Japan (1). In 2011, a large outbreak due to the consumption of raw beef was reported. A total of 181 patients, including 21 patients with acute encephalopathy (AE) and five deaths among 34 patients with hemolytic-uremic syndrome (HUS), were recorded (2, 3). EHEC O111:H8 and O157:H7 were isolated from the patients; however, the Shiga toxin 2a gene (*stx*<sub>2</sub>)-positive EHEC O111:H8 strain was the primary cause of the severe infection, according to serological testing of the patients (4). To reveal the characteristics of the strain, we determined its complete genome sequence, as the complete genome of the *stx*<sub>2</sub>-positive EHEC O111 strain has not been reported.

EHEC O111:H8 110512 was isolated as one of a few colonies grown on deoxycholate-hydrogen sulfide-lactose agar (Nissui Pharmaceutical Co., Ltd., Tokyo, Japan) from the fecal sample of a 6-year-old boy with HUS and AE. The strain was stored in 0.5% NaCl casein peptone medium (Kohjin Bio Co., Ltd., Saitama, Japan) until use. For DNA extraction, a single colony was inoculated into buffered peptone water (Nissui). The overnight culture was subjected to genomic DNA extraction by a DNeasy blood and tissue kit (Qiagen, Venlo, Netherlands). The genomic DNA libraries were prepared using a Nextera DNA sample prep kit (Epicentre Biotechnologies, Madison, WI, USA), followed by paired-end sequencing (2 × 150-mer) with a MiSeq instrument (Illumina, San Diego, CA). Additionally, the strain was sequenced on a PacBio RS II instrument (Pacific Biosciences, Menlo Park, CA). The genomic DNA was fragmented prior to PacBio sequencing using the Megaruptor (Diagenode, Denville, NJ). PacBio sequencing was performed by using the SMRTbell template prep kit 1.0 and polymerase binding kit P6 after size selection using BluePippin (Sage Science, Beverly, MA) with a cutoff value of 20 kb. In bioinformatic analysis, default parameters were used for all software. The complete genome of the strain was obtained by using both PacBio and Illumina reads. PacBio reads were assembled using the Hierarchical Genome Assembly Process version 3 in SMRT Analysis software (5). The minimum seed length for the assembly was 6,000. Short plasmids (≤10 kb) were assembled using A5 MiSeq software version 20140604 with Illumina short reads (6). The tentative circular contigs were subjected to error correction in three iterations using Pilon version 1.18 with default parameters and

**Citation** Sekizuka T, Lee K, Kimata K, Isobe J, Kuroda M, Iyoda S, Ohnishi M, Sata T, Watahiki M. 2019. Complete genome sequence of an enterohemorrhagic *Escherichia coli* O111:H8 strain recovered from a large outbreak in Japan associated with consumption of raw beef. *Microbiol Resour Announc* 8:e00882-19. <https://doi.org/10.1128/MRA.00882-19>.

**Editor** David A. Baltrus, University of Arizona

**Copyright** © 2019 Sekizuka et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Masanori Watahiki, [masanori.watahiki01@pref.toyama.lg.jp](mailto:masanori.watahiki01@pref.toyama.lg.jp).

**Received** 29 July 2019

**Accepted** 30 August 2019

**Published** 3 October 2019

**TABLE 1** Statistics of the complete genome of the enterohemorrhagic *Escherichia coli* O111:H8 strain 110512

Statistic	Data for <sup>a</sup> :							
	Chromosome	pO111-110512_1	pO111-110512_2	pO111-110512_3	pO111-110512_4	pO111-110512_5	pO111-110512_6	pO111-110512_7
Size (kb)	5,302,257	118,827	78,470	46,350	6,795	6,674	5,423	1,546
No. of CDSs <sup>b</sup>	5,119	131	81	57	6	9	6	1
No. of rRNAs	22	0	0	0	0	0	0	0
No. of tRNAs	107	0	0	0	0	0	0	0
No. of prophages	13	0	0	0	0	0	0	0
GC content (%)	50.6	53.3	50.0	40.7	50.8	50.3	47.2	51.7
Chromosomal sequence type or plasmid Inc type	ST16	IncQ1, IncP	IncFII (pRSB107)	IncX1	ND	ND	ND	Col (MG828)
Antimicrobial resistance gene(s)	<i>bla</i> <sub>EC-18</sub>	<i>bla</i> <sub>TEM-1Br</sub> , <i>strB</i> , <i>strA</i> , <i>sul2</i>	ND	ND	ND	ND	ND	ND
Accession no.	AP019761	AP019762	AP019763	AP019764	AP019765	AP019766	AP019767	AP019768

<sup>a</sup> ND, not detected.<sup>b</sup> CDSs, coding DNA sequences.

Illumina short reads (7). Annotation was performed in Prokka version 1.11 (8), InterPro version 49.0 (9), and NCBI BLASTP/BLASTX.

In Illumina and PacBio sequencing, the read numbers were 5,894,924 and 195,083 (mean read length, 7,798 bp), respectively. The total read lengths for Illumina and PacBio sequencing were 732,012,018 bp and 1,521,220,270 bp, respectively; the mean coverage was 132-fold and 273-fold, respectively. The complete genome sequence consists of a chromosome and seven plasmids as shown in Table 1. The strain carried a novel Stx2 prophage but did not carry a Stx1 prophage. These data would be helpful for the characterization of *stx*<sub>2</sub>-positive EHEC O111 and improve the resolution of the surveillance using whole-genome sequencing (10, 11).

**Data availability.** The complete genome sequence of EHEC O111:H8 strain 110512 was submitted to DDBJ/ENA/GenBank under the accession numbers AP019761 through AP019768. The primary data were deposited in the NCBI primary data archive, SRA, under the reference number DRP005302.

## ACKNOWLEDGMENTS

This work was partially supported by a special research grant (H23-TOKUBETU-SHITEI-004) and a Health Labour Sciences research grant (H24-SHINKO-IPPAN-012) from the Ministry of Health, Labour, and Welfare (Japan) and by the Research Program on Emerging and Re-emerging Infectious Diseases from the Japan Agency for Medical Research and Development (AMED) (grant number JP18fk0108065).

## REFERENCES

- National Institute of Infectious Diseases. 2018. Enterohemorrhagic *Escherichia coli* (EHEC) infection, as of March 2018, Japan. Infectious Agents Surveillance Report, vol 39, no. 5. National Institute of Infectious Diseases, Tokyo, Japan. <https://www.niid.go.jp/niid/en/iasr-vol39-e/865-iasr/8042-459te.html>.
- Yahata Y, Misaki T, Ishida Y, Nagira M, Watahiki M, Isobe J, Terajima J, Iyoda S, Mitobe J, Ohnishi M, Sata T, Taniguchi K, Tada Y, Okabe N, E. coli O111 Outbreak Investigation Team. 2015. Epidemiological analysis of a large enterohaemorrhagic *Escherichia coli* O111 outbreak in Japan associated with haemolytic uraemic syndrome and acute encephalopathy. *Epidemiol Infect* 143:2721–2732. <https://doi.org/10.1017/S0950268814003641>.
- Watahiki M, Isobe J, Kimata K, Shima T, Kanatani J, Shimizu M, Nagata A, Kawakami K, Yamada M, Izumiya H, Iyoda S, Morita-Ishihara T, Mitobe J, Terajima J, Ohnishi M, Sata T. 2014. Characterization of enterohemorrhagic *Escherichia coli* O111 and O157 strains isolated from outbreak patients in Japan. *J Clin Microbiol* 52:2757–2763. <https://doi.org/10.1128/JCM.00420-14>.
- Isobe J, Shima T, Kanatani J, Kimata K, Shimizu M, Kobayashi N, Tanaka T, Iyoda S, Ohnishi M, Sata T, Watahiki M. 2014. Serodiagnosis using microagglutination assay during the food-poisoning outbreak in Japan caused by consumption of raw beef contaminated with enterohemorrhagic *Escherichia coli* O111 and O157. *J Clin Microbiol* 52:1112–1118. <https://doi.org/10.1128/JCM.03469-13>.
- 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>.
- Coil D, Jospin G, Darling AE. 2015. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics* 31:587–589. <https://doi.org/10.1093/bioinformatics/btu661>.
- 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>.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
- Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang

- HY, Dosztanyi Z, El-Gebali S, Fraser M, Gough J, Haft D, Holliday GL, Huang H, Huang X, Letunic I, Lopez R, Lu S, Marchler-Bauer A, Mi H, Mistry J, Natale DA, Necci M, Nuka G, Orengo CA, Park Y, Pesseat S, Piovesan D, Potter SC, Rawlings ND, Redaschi N, Richardson L, Rivoire C, Sangrador-Vegas A, Sigrist C, Sillitoe I, Smithers B, Squizzato S, Sutton G, Thanki N, Thomas PD, Tosatto SC, Wu CH, Xenarios I, Yeh LS, Young SY, Mitchell AL. 2017. InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res* 45:D190–D199. <https://doi.org/10.1093/nar/gkw1107>.
10. Lee KI, Morita-Ishihara T, Iyoda S, Ogura Y, Hayashi T, Sekizuka T, Kuroda M, Ohnishi M, EHEC Working Group in Japan. 2017. A geographically widespread outbreak investigation and development of a rapid screening method using whole genome sequences of enterohemorrhagic *Escherichia coli* O121. *Front Microbiol* 8:701. <https://doi.org/10.3389/fmicb.2017.00701>.
11. Dallman TJ, Byrne L, Ashton PM, Cowley LA, Perry NT, Adak G, Petrovska L, Ellis RJ, Elson R, Underwood A, Green J, Hanage WP, Jenkins C, Grant K, Wain J. 2015. Whole-genome sequencing for national surveillance of Shiga toxin-producing *Escherichia coli* O157. *Clin Infect Dis* 61:305–312. <https://doi.org/10.1093/cid/civ318>.