

Whole-Genome Sequencing Applied to the Molecular Epidemiology of Shiga Toxin-Producing *Escherichia coli* O157:H7 in Argentina

Claudia Carolina Carbonari,^a Nahuel Fittipaldi,^{b,c} Sarah Teatero,^b Taryn B. T. Athey,^b Luis Pianciola,^d Marcelo Masana,^{e†} Roberto G. Melano,^{b,c} Marta Rivas,^a Isabel Chinen^a

Servicio Fisiopatogenia, Departamento Bacteriología, Instituto Nacional de Enfermedades Infecciosas-ANLIS "Dr. Carlos G. Malbrán," Buenos Aires, Argentina^a; Public Health Ontario, Toronto Laboratories, Toronto, Ontario, Canada^b; Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada^c; Subsecretaría de Salud, Laboratorio Central, Neuquén, Argentina^d; Instituto Tecnología de Alimentos, Centro de Investigación de Agroindustria, Instituto Nacional de Tecnología Agropecuaria (INTA), Morón, Argentina^e

† Deceased.

Shiga toxin-producing *Escherichia coli* strains are worldwide associated with sporadic human infections and outbreaks. In this work, we report the availability of high-quality draft whole-genome sequences for 19 O157:H7 strains isolated in Argentina.

Received 6 October 2016 Accepted 7 October 2016 Published 1 December 2016

Citation Carbonari CC, Fittipaldi N, Teatero S, Athey TBT, Pianciola L, Masana M, Melano RG, Rivas M, Chinen I. 2016. Whole-genome sequencing applied to the molecular epidemiology of Shiga toxin-producing *Escherichia coli* O157:H7 in Argentina. *Genome Announc* 4(6):e01341-16. doi:10.1128/genomeA.01341-16.

Copyright © 2016 Carbonari 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 Claudia Carolina Carbonari, ccarbonari@anlis.gov.ar.

Shiga toxin-producing *Escherichia coli* (STEC) strains are a common cause of both sporadic infections and food and waterborne outbreaks. In Argentina, hemolytic uremic syndrome (HUS) is endemic with approximately 400 new cases reported annually.

STEC O157:H7 is the dominant serotype and *stx*_{2a}/*stx*_{2c} the prevalent genotype. The main source of O157 is cattle. Previous studies have shown that O157:H7 strains belonging to clade 8 and lineage I/II by lineage specific polymorphism assay (LSPA-6) are prevalent in human and cattle isolates. Also, it has been demonstrated that *stx*_{2a} strains belonging to clade 8 were high toxin producers (1, 2).

This work reports the availability of high-quality draft whole-genome sequences for 19 (11 human and eight bovine) O157:H7 strains isolated in Argentina. The isolates had previously been characterized phenotypically and genotypically by PFGE, single nucleotide polymorphism (SNP), and LSPA-6 (3, 4).

Strains were grown overnight on Tryptic soy agar (BD-Difco, Le Pont de Claix, France), and genomic DNA extraction (QIAamp DNA Mini Kit, Qiagen Group) and quantification (Qubit fluorometer, Invitrogen, Eugene, Oregon, USA) were performed according to the manufacturer protocols. Libraries were prepared using the Nextera XT DNA library preparation kit (Illumina, San Diego, CA, USA), and sequenced as paired-end reads (150 bp +

TABLE 1 Accession numbers and assembly metrics of the annotated STEC O157:H7 strains

Isolate no.	Origin	City or province/yr of isolation	NCBI accession no.	No. of scaffolds	Genome size (bp)	<i>N</i> ₅₀	G+C content (%)
GN1209	<i>Homo sapiens</i>	La Pampa/2006	LXJJ00000000	112	5,322,179	201,363	50
GN1210	<i>Homo sapiens</i>	Buenos Aires/2008	LXJI00000000	117	5,431,510	267,470	50
GN1211	<i>Homo sapiens</i>	Buenos Aires/2007	LXJH00000000	117	5,303,086	159,713	51
GN1212	<i>Bos taurus</i>	Buenos Aires/2007	LXJG00000000	126	5,320,347	185,674	50
GN1213	<i>Homo sapiens</i>	Buenos Aires/2006	LXJF00000000	143	5,349,522	170,212	50
GN1214	<i>Bos taurus</i>	Buenos Aires/2007	LXJE00000000	162	5,361,881	243,406	50
GN1215	<i>Homo sapiens</i>	Buenos Aires/2007	LXJD00000000	151	5,427,792	171,358	51
GN1216	<i>Bos taurus</i>	Buenos Aires/2008	LXJC00000000	137	5,378,586	237,563	50
GN1217	<i>Homo sapiens</i>	Chubut/2006	LXJB00000000	158	5,476,251	200,755	50
GN1218	<i>Bos taurus</i>	Santa Fe/2007	LXJA00000000	170	5,473,683	179,381	50
GN1219	<i>Homo sapiens</i>	Chubut/2007	LXIZ00000000	123	5,428,694	200,826	51
GN1220	<i>Bos taurus</i>	Buenos Aires/2007	LXIY00000000	135	5,441,312	180,736	50
GN1221	<i>Homo sapiens</i>	Buenos Aires/2002	LXIX00000000	114	5,328,396	237,941	50
GN1222	<i>Bos taurus</i>	Santa Fe/2007	LXIW00000000	140	5,334,541	170,808	50
GN1223	<i>Homo sapiens</i>	Santa Fe/2008	LXIV00000000	105	5,300,977	248,557	50
GN1224	<i>Bos taurus</i>	Buenos Aires/2006	LXIU00000000	116	5,238,450	196,625	50
GN1225	<i>Homo sapiens</i>	Buenos Aires/2007	LXIT00000000	145	5,417,114	201,254	50
GN1226	<i>Homo sapiens</i>	Rio Negro/2007	LXIS00000000	116	5,298,547	171,490	50
GN1227	<i>Bos taurus</i>	Corrientes/2007	LXIR00000000	132	5,312,947	201,355	50

150 bp) on an Illumina MiSeq instrument (Illumina 1.9, San Diego, CA, USA). Quality of raw reads was assessed with FastQC v0.11.5 (5) and custom scripts. Paired-end reads were then *de novo* assembled using the A5 pipeline (6) and obtained contigs reordered against reference strain TW14359 (3) using Mauve v 2.3.1 (7). The sequences were annotated with the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (8).

Further analysis of the draft genomes will be included in a future publication.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the BioProject number PRJNA299801. The versions described in this paper are the first draft genome sequences for the 19 STEC O157:H7 strains, under the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

We thank the staff at the Public Health Ontario Genome Core.

This work was supported by the Ministry of Health of Argentina and by Public Health Ontario.

REFERENCES

- Mellor GE, Sim EM, Barlow RS, D'Astek BA, Galli L, Chinen I, Rivas M, Gobius KS. 2012. Phylogenetically related Argentinian and Australian *Escherichia coli* O157:H7 isolates are distinguished by virulence clades and alternative Shiga toxin 1 and 2 prophages. *Appl Environ Microbiol* 78:4724–4731.
- Pianciola L, D'Astek BA, Mazzeo M, Chinen I, Masana M, Rivas M. 2016. Genetic features of human and bovine *Escherichia coli* O157:H7 strains isolated in Argentina. *Int J Med Microbiol* 306:123–130. <http://dx.doi.org/10.1016/j.ijmm.2016.02.005>.
- Manning SD, Motiwala AS, Springman AC, Qi W, Lacher DW, Ouellette LM, Mladonicky JM, Somsel P, Rudrik JT, Dietrich SE, Zhang W, Swaminathan B, Alland D, Whittam TS. 2008. Variation in virulence among clades of *Escherichia coli* O157:H7 associated with disease outbreaks. *Proc Natl Acad Sci U S A* 105:4868–4873.
- Riordan JT, Viswanath SB, Manning SD, Whittam TS. 2008. Genetic differentiation of *Escherichia coli* O157:H7 clades associated with human disease by real-time PCR. *J Clin Microbiol* 46:2070–2073.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Tritt A, Eisen JA, Facciotti MT, Darling AE. 2012. An integrated pipeline for *de novo* assembly of microbial genomes. *PLoS One* 7:e42304. <http://dx.doi.org/10.1371/journal.pone.0042304>.
- Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. <http://dx.doi.org/10.1101/gr.2289704>.
- Klimke W, Agarwala R, Badretidin A, Chetvernin S, Ciuffo S, Fedorov B, Kiryutin B, O'Neill K, Resch W, Resenchuk S, Schafer S, Tolstoy I, Tatusova T. 2009. The National Center for Biotechnology Information's protein clusters database. *Nucleic Acids Res* 37:D216–D223. <http://dx.doi.org/10.1093/nar/gkn734>.