CrossMark
click for updates

Draft Genome Sequences of *Escherichia coli* O157:H7 Strains Rafaela_II (Clade 8) and 7.1_Anguil (Clade 6) from Cattle in Argentina

 Ariel Fernando Amadio,a,c Natalia Amigo,b,c Andrea Fabiana Puebla,b Marisa Diana Farber,b,c Angel Adrián Cataldib,c

Estación Experimental Agropecuaria Rafaela, Instituto Nacional de Tecnología Agropecuaria (INTA), Rafaela, Santa Fe, Argentinaa; Centro de Investigación en Ciencias Veterinarias y Agronómicas (CICVyA), Instituto de Biotecnología, Instituto Nacional de Tecnología Agropecuaria (INTA), Hurlingham, Buenos Aires, Argentinab; Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentinac

Escherichia coli **O157:H7 is a major etiologic agent of diseases in humans that cause diarrhea, hemorrhagic colitis, and hemolytic-uremic syndrome. Here, we report the draft genome sequences of two strains isolated from cattle that had high levels of Shiga toxin 2 and high lethality in mice.**

Address correspondence to Angel Adrián Cataldi, cataldi.angeladrian@inta.gob.ar.

*E*scherichia coli O157:H7 is the etiologic agent of hemolytic-uremic syndrome. This is the leading cause of chronic renal failure in children in Argentina and several other countries (1). This bacterium produces Shiga toxin (Stx) types 1 and 2 (2, 3), which are responsible for systemic damage. The main reservoir of *E. coli* O157:H7 is cattle, which harbor this organism in their intestinal tract (4, 5), especially on the lymphoid follicle-dense mucosa at the terminal rectum (6). Fecal contamination of meat during slaughter, the use of feces as fertilizer, and the contamination of drinking water (5, 7) cause the entry of these bacteria into the human food chain. Single nucleotide polymorphism (SNP) typing was previously used to define nine *E. coli* O157:H7 clades, and the clade 8 strains were found to be more associated with severe disease (8).

This work reports the draft genome sequences of two *E. coli* O157:H7 strains isolated from cattle in the central humid Pampas, Argentina, in 2009. Strain Rafaela_II belongs to clade 8, and strain 7.1_Anguil belongs to clade 6. Both strains produced elevated levels of Shiga toxin 2 and had high lethality in mice (N. Amigo, E. Mercado, A. Bentancor, P. Singh, D. Vilte, E. Gerhardt, E. Zotta, C. Ibarra, S. D. Manning, M. Larzábal, and A. Cataldi, unpublished data).

Genomic DNA was isolated using a standard chloroform-isoamilic alcohol extraction. Paired-end Nextera XT libraries were constructed (500-bp insert size) and sequenced in an Illumina MiSeq sequencer (2 × 250 bp). Raw sequences were quality trimmed with Sickle (9), resulting in 1,262,211 and 1,048,969 sequences for 7.1_Anguil and Rafaela_II, respectively. *De novo* assembly was done using SPAdes version 3.1.0 (10). For 7.1_Anguil, 244 scaffolds >500 bp were obtained, the largest being 375,850 bp, with an N_{50} of 146,432 bp and a G+C content of 50.46%. For Rafaela_II, 232 scaffolds >500 bp were obtained, the largest being 375,042 bp, with an N_{50} of 147,588 bp and a G+C content of 50.28%. Scaffolds were ordered using ABACAS (11) and compared with the genomes of *E. coli* O157:H7 strains Tw14359 (12)

and EDL933 (13) using BLAST and ACT (14) to analyze structural differences. Variants at the nucleotide level were analyzed using *breseq* (15) and the same references. Both strains carry the large virulence plasmid pO157, but strain 7.1_Anguil has an additional plasmid highly homologous to avian pathogenic *E. coli* (APEC) strain 7122 (O78:K80:H9) plasmid pChi7122-3.

Genes were predicted with Prodigal version 2.6.1 (16), obtaining 5,643 and 5,439 genes for 7.1_Anguil and Rafaela_II, respectively. Using GET_HOMOLOGUES (17), we compared the gene content using the OrthoMCL methodology (18) to obtain shared and unique genes using strains TW14359 and EDL933 as references. The comparison of shared genes showed a closer phylogenetic relationship between isolate 7.1_Anguil and reference strain EDL933, while isolate Rafaela_II is closely related to TW14359. This is also supported by a previously performed SNP analysis (15). Based on core and pangenome analysis, a considerable number of exclusive genes from Rafaela_II (272) and 7.1_Anguil (447) are under analysis.

Nucleotide sequence accession numbers. These whole-genome shotgun (WGS) projects have been deposited at DDBJ/EMBL/GenBank under the accession numbers LAZD00000000 for 7.1_Anguil and LAYW00000000 for Rafaela_II. The versions described in this paper are LAZD01000000 and LAYW01000000, respectively.

REFERENCES

1. **Karch H, Tarr PI, Bielaszewska M**. 2005. Enterohaemorrhagic *Escherichia coli* in human medicine. Int J Med Microbiol **295:**405–418. http://dx.doi.org/10.1016/j.ijmm.2005.06.009.
2. **Jackson MP, Neill RJ, O'Brien AD, Holmes RK, Newland JW**. 1987. Nucleotide sequence analysis and comparison of the structural genes for Shiga-like toxin I and Shiga-like toxin II encoded by bacteriophages from

Escherichia coli 933. FEMS Microbiol Lett **44**:109–114. http://dx.doi.org/10.1111/j.1574-6968.1987.tb02252.x.

3. Tesh VL, O'Brien AD. 1991. The pathogenic mechanisms of Shiga toxin and the Shiga-like toxins. Mol Microbiol **5**:1817–1822. http://dx.doi.org/10.1111/j.1365-2958.1991.tb00805.x.

4. Meichtri L, Miliwebsky E, Gioffré A, Chinen I, Baschkier A, Chillemi G, Guth BE, Masana MO, Cataldi A, Rodríguez HR, Rivas M. 2004. Shiga toxin-producing *Escherichia coli* in healthy young beef steers from Argentina: prevalence and virulence properties. Int J Food Microbiol **96**: 189–198. http://dx.doi.org/10.1016/j.ijfoodmicro.2004.03.018.

5. Hussein HS, Bollinger LM. 2005. Prevalence of Shiga toxin-producing *Escherichia coli* in beef cattle. J Food Prot **68**:2224–2241.

6. Grauke LJ, Kudva IT, Yoon JW, Hunt CW, Williams CJ, Hovde CJ. 2002. Gastrointestinal tract location of *Escherichia coli* O157:H7 in ruminants. Appl Environ Microbiol **68**:2269–2277. http://dx.doi.org/10.1128/AEM.68.5.2269-2277.2002.

7. Olsen SJ, Miller G, Breuer T, Kennedy M, Higgins C, Walford J, McKee G, Fox K, Bibb W, Mead P. 2002. A waterborne outbreak of *Escherichia coli* O157:H7 infections and hemolytic uremic syndrome: implications for rural water systems. Emerg Infect Dis **8**:370–375. http://dx.doi.org/10.3201/eid0804.000218.

8. 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. http://dx.doi.org/10.1073/pnas.0710834105.

9. Joshi NA, Fass JN. 2011. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (version 1.33). https://github.com/najoshi/sickle.

10. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin 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. http://dx.doi.org/10.1089/cmb.2012.0021.

11. Assefa S, Keane TM, Otto TD, Newbold C, Berriman M. 2009. ABACAS: algorithm-based automatic contiguation of assembled sequences. Bioinformatics **25**:1968–1969. http://dx.doi.org/10.1093/bioinformatics/btp347.

12. Kulasekara BR, Jacobs M, Zhou Y, Wu Z, Sims E, Saenphimmachak C, Rohmer L, Ritchie JM, Radey M, McKevitt M, Freeman TL, Hayden H, Haugen E, Gillett W, Fong C, Chang J, Beskhlebnaya V, Waldor MK, Samadpour M, Whittam TS, Kaul R, Brittnacher M, Miller SI. 2009. Analysis of the genome of the *Escherichia coli* O157:H7 2006 spinach-associated outbreak isolate indicates candidate genes that may enhance virulence. Infect Immun **77**:3713–3721. http://dx.doi.org/10.1128/IAI.00198-09.

13. Latif H, Li HJ, Charusanti P, Palsson BØ, Aziz RK. 2014. A gapless, unambiguous genome sequence of the enterohemorrhagic *Escherichia coli* O157:H7 strain EDL933. Genome Announc **2**(4):e00821-14. http://dx.doi.org/10.1128/genomeA.00821-14.

14. Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. 2005. ACT: the Artemis comparison tool. Bioinformatics **21**: 3422–3423. http://dx.doi.org/10.1093/bioinformatics/bti553.

15. Barrick JE, Colburn G, Deatherage DE, Traverse CC, Strand MD, Borges JJ, Knoester DB, Reba A, Meyer AG. 2014. Identifying structural variation in haploid microbial genomes from short-read resequencing data using *breseq*. BMC Genomics **15**:1039. http://dx.doi.org/10.1186/1471-2164-15-1039.

16. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics **11**:119. http://dx.doi.org/10.1186/1471-2105-11-119.

17. Contreras-Moreira B, Vinuesa P. 2013. GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. Appl Environ Microbiol **79**:7696–7701. http://dx.doi.org/10.1128/AEM.02411-13.

18. Li L, Stoeckert CJ, Jr, Roos DS. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res **13**:2178–2189. http://dx.doi.org/10.1101/gr.1224503.