

Genome Sequence of Enterotoxigenic *Escherichia coli* Strain B2C

T. P. Vipin Madhavan,^a Jason A. Steen,^b Philip Hugenholtz,^b Harry Sakellaris^a

School of Medical Sciences, Griffith Health Centre, Griffith University, Gold Coast Campus, Queensland, Australia^a; Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, University of Queensland, St. Lucia, Queensland, Australia^b

T.P.V.M. and J.A.S. contributed equally to this work.

Enterotoxigenic *Escherichia coli* (ETEC) is a major cause of diarrheal disease around the globe, causing an estimated 380,000 deaths annually. The disease is caused by a wide variety of strains. Here, we report the genome sequence of ETEC strain B2C, which was isolated from an American soldier in Vietnam.

Received 20 March 2014 Accepted 24 March 2014 Published 10 April 2014

Citation Vipin Madhavan TP, Steen JA, Hugenholtz P, Sakellaris H. 2014. Genome sequence of enterotoxigenic *Escherichia coli* strain B2C. *Genome Announc.* 2(2):e00247-14. doi:10.1128/genomeA.00247-14.

Copyright © 2014 Vipin Madhavan et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Harry Sakellaris, h.sakellaris@griffith.edu.au.

Enterotoxigenic *Escherichia coli* (ETEC) causes an estimated 200 million cases of watery diarrhea and up to 380,000 deaths annually (1, 2). The major burden of disease is in developing countries lacking adequate water treatment facilities (3). The severity and symptoms of the disease can vary with the strain of ETEC causing it (4). The genomes of only two ETEC strains that infect humans, *E. coli* H10407 and E24377A, have been completely sequenced (5, 6). Strain H10407 expresses the archetypal ETEC pilus colonization factor, CFA/I, and a variety of established and putative virulence factors. E24377A expresses the colonization factors CS1 and CS3 and shares most of the same virulence genes found in H10407 (5). A draft genome of ETEC strain B7A, which expresses the CS6 colonization factor, is also available (7). B2C is an O6:H16 serotype strain of ETEC isolated in 1971 from a soldier in Vietnam (8). Since then, it has become an important laboratory strain that has been used in a variety of studies (9–13). Unlike the sequenced ETEC strains, B2C expresses CS2 and CS3 pili (9).

Genomic DNA was extracted from B2C with the Fermentas GeneJET genomic DNA purification kit (Thermo Fisher Scientific, Australia), according to the manufacturer's protocol. One hundred nanograms of genomic DNA was used to generate a library using the Nextera DNA sample prep kit (Illumina, CA). The libraries were sequenced at the Queensland Centre for Medical Genomics, University of Queensland, Queensland (QLD), Australia, on an Illumina HiSeq 2000 sequencing system (Illumina, CA). Sequencing generated approximately 5.7 million read pairs for a ~200× coverage of the genome. The read pairs were overlapped where possible using the SeqPrep software (<https://github.com/jstjohn/SeqPrep>) and trimmed for quality using the Nsoni software (<http://www.vicbioinformatics.com/software/nesoni.shtml>). A *de novo* assembly of the overlapped and quality trimmed reads was generated using Velvet version 1.2.07 (14). The final assembly consists of 313 contigs with a total size of 5,018,127 bp. Further assembly in order to cover the entire genome was difficult due to the presence of a large number of repeat sequences and insertion elements. The size of the genome of ETEC B2C is estimated to be 5.2 Mb. The draft genome was annotated

using the NCBI Prokaryotic Genome Annotation Pipeline (http://www.ncbi.nlm.nih.gov/genome/annotation_prok/) and manually curated using Artemis (15). An analysis of the genome showed that ETEC B2C has 4,990 genes, 4,804 coding sequences (CDS), 94 pseudogenes, 2 clustered regularly interspaced short palindromic repeat (CRISPR) arrays, 9 rRNAs, 83 tRNAs, and 64 frameshifted genes. Classical ETEC virulence genes, including pilus colonization factors of the chaperone usher family, and enterotoxins were also found. The genes encoding colonization factors include two operons encoding CS2 and CS3 pili and the AraC family transcriptional activator, *rms*, which control both CS2 and CS3 expression. The genome also carries *eltA* and *eltB*, the genes encoding the two subunits of heat-labile enterotoxin, and *sta2* and *east1*, both of which encode heat-stable enterotoxins. Since these virulence genes are carried by plasmids in other ETEC strains, we presume that they are also carried on plasmids in B2C.

Nucleotide sequence accession number. The draft genome sequence of enterotoxigenic *E. coli* (ETEC) O6:H16:CFA/II strain B2C was deposited in Genbank with accession no. AUZS00000000.

ACKNOWLEDGMENTS

T.P.V.M. is supported by a Griffith University International Postgraduate Research Scholarship (GUIPRS) and a Griffith University Postgraduate Research Scholarship (GUPRS).

REFERENCES

1. Steffen R, Castelli F, Dieter Nothdurft H, Rombo L, Jane Zuckerman N. 2005. Vaccination against enterotoxigenic *Escherichia coli*, a cause of travelers' diarrhea. *J. Travel Med.* 12:102–107. <http://dx.doi.org/10.2310/706.0.2005.12207>.
2. Wennerås C, Erling V. 2004. Prevalence of enterotoxigenic *Escherichia coli*-associated diarrhoea and carrier state in the developing world. *J. Health Popul. Nutr.* 22:370–382.
3. Fleckenstein JM, Hardwidge PR, Munson GP, Rasko DA, Sommerfelt H, Steinsland H. 2010. Molecular mechanisms of enterotoxigenic *Escherichia coli* infection. *Microbes Infect.* 12:89–98. <http://dx.doi.org/10.1016/j.micinf.2009.10.002>.
4. Porter CK, Riddle MS, Tribble DR, Louis Bougeois A, McKenzie R, Isidean SD, Sebeny P, Savarino SJ. 2011. A systematic review of experi-

- mental infections with enterotoxigenic *Escherichia coli* (ETEC). *Vaccine* 29:5869–5885. <http://dx.doi.org/10.1016/j.vaccine.2011.05.021>.
5. Crossman LC, Chaudhuri RR, Beatson SA, Wells TJ, Desvaux M, Cunningham AF, Petty NK, Mahon V, Brinkley C, Hobman JL, Savarino SJ, Turner SM, Pallen MJ, Penn CW, Parkhill J, Turner AK, Johnson TJ, Thomson NR, Smith SG, Henderson IR. 2010. A commensal gone bad: complete genome sequence of the prototypical enterotoxigenic *Escherichia coli* strain H10407. *J. Bacteriol.* 192:5822–5831. <http://dx.doi.org/10.1128/JB.00710-10>.
 6. Sahl JW, Rasko DA. 2012. Analysis of global transcriptional profiles of enterotoxigenic *Escherichia coli* isolate E24377A. *Infect. Immun.* 80:1232–1242. <http://dx.doi.org/10.1128/IAI.06138-11>.
 7. Rasko DA, Rosovitz MJ, Myers GS, Mongodin EF, Fricke WF, Gajer P, Crabtree J, Sebaihia M, Thomson NR, Chaudhuri R, Henderson IR, Sperandio V, Ravel J. 2008. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J. Bacteriol.* 190:6881–6893. <http://dx.doi.org/10.1128/JB.00619-08>.
 8. DuPont HL, Formal SB, Hornick RB, Snyder MJ, Libonati JP, Sheahan DG, LaBrec EH, Kalas JP. 1971. Pathogenesis of *Escherichia coli* diarrhea. *N. Engl. J. Med.* 285:1–9. <http://dx.doi.org/10.1056/NEJM197107012850101>.
 9. Knutton S, Lloyd DR, Candy DC, McNeish AS. 1984. Ultrastructural study of adhesion of enterotoxigenic *Escherichia coli* to erythrocytes and human intestinal epithelial cells. *Infect. Immun.* 44:519–527.
 10. Knutton S, Lloyd DR, Candy DC, McNeish AS. 1984. In vitro adhesion of enterotoxigenic *Escherichia coli* to human intestinal epithelial cells from mucosal biopsies. *Infect. Immun.* 44:514–518.
 11. Duchet-Suchaux M, Menanteau P, van Zijderveld FG. 1992. Passive protection of suckling infant mice against F41-positive enterotoxigenic *Escherichia coli* strains by intravenous inoculation of the dams with monoclonal antibodies against F41. *Infect. Immun.* 60:2828–2834.
 12. Qadri F, Girón JA, Helander A, Begum YA, Asaduzzaman M, Xicohténcatl-Cortes J, Negrete E, Albert MJ. 2000. Human antibody response to longus type IV pilus and study of its prevalence among enterotoxigenic *Escherichia coli* in Bangladesh by using monoclonal antibodies. *J. Infect. Dis.* 181:2071–2074. <http://dx.doi.org/10.1086/315507>.
 13. Isidean SD, Riddle MS, Savarino SJ, Porter CK. 2011. A systematic review of ETEC epidemiology focusing on colonization factor and toxin expression. *Vaccine* 29:6167–6178. <http://dx.doi.org/10.1016/j.vaccine.2011.06.084>.
 14. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 18:821–829. <http://dx.doi.org/10.1101/gr.074492.107>.
 15. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16:944–945. <http://dx.doi.org/10.1093/bioinformatics/16.10.944>.