genomeA
Journals.ASM.org

Draft Genome Sequence of *Clostridium butyricum* Strain NOR 33234, Isolated from an Elderly Patient with Diarrhea

Jamie S. L. Kwok,[a] Margaret Ip,[b] Ting-Fung Chan,[c] Wai-Yip Lam,[a] Stephen K. W. Tsui[a,d]

School of Biomedical Sciences,[a] Department of Microbiology,[b] School of Life Sciences,[c] and Centre for Microbial Genomics and Proteomics,[d] The Chinese University of Hong Kong, Hong Kong, China

***Clostridium butyricum* is one of the species frequently present in patients' stool samples. However, the identification of this species is sometimes difficult. Here, we present the draft genome of *Clostridium butyricum* NOR 33234, which was isolated from a patient with suspected *Clostridium difficile* infection-associated diarrhea and resembles *Clostridium clostridioforme* in biochemical tests.**

Received 14 November 2014 **Accepted** 20 November 2014 **Published** 24 December 2014

Citation Kwok JSL, Ip M, Chan T-F, Lam W-Y, Tsui SKW. 2014. Draft genome sequence of *Clostridium butyricum* strain NOR 33234, isolated from an elderly patient with diarrhea. Genome Announc. 2(6):e01356-14. doi:10.1128/genomeA.01356-14.

Address correspondence to Stephen K. W. Tsui, kwtsui@cuhk.edu.hk.

The varied roles of *Clostridium butyricum* range from a promising biofuel producer (1) to a neurotoxigenic pathogen (2). Toxin-free strains are used in probiotics, especially in preventing *Clostridium diffiicle*-related diarrhea (3). Recently, the genome of strain 5521 carrying botulinum neurotoxin type E (*boNT/E*) has been sequenced (4).

A stool sample from a 94-year-old male patient with suspected *C. diffiicle* infection (CDI)-associated diarrhea was collected. A strain was isolated and identified by VITEK2 as *Clostridium clostridioforme*. Among comprehensive biochemical tests, only α- and β-galactosidase activities were positive ($^+$) (AGALi$^+$, BGALi$^+$). Previous reports suggest that besides *C. difficile*, *Clostridium perfringens* (5, 6), or other *Clostridium* spp. may also cause antibiotic-associated diarrhea, of which CDI is the major cause. To understand the nature of the strain, we sequenced the genome using Ion Torrent single-end libraries, of which 6,595,984 reads were assembled into 70 long contigs (>10,000 bp) using MIRA (version 4.0.2) (7) and gap5 (8), with an N_{50} of 156,811 bp. In contrast to biochemical identification, 16S analysis using 16SpathDB 2.0 (9) identified our sample as *C. butyricum* with 99.0% nucleotide identity, within the species definition threshold. The GC content of the genome is 28.37%, which is also similar to that of published *C. butyricum* genomes and very different from that of *C. clostridioforme* (GC content, ~49%).

Protein prediction is done using the NCBI prokaryotic genome annotation pipeline (10), resulting in 4,357 genes, with 4,085 coding genes and 272 pseudogenes. When all the coding genes were searched against nonredundant protein databases in NCBI, 3,901 genes had significant hits (*E* value $\leq 1e^{-30}$). Among them, 3,714 (95.21%) have top hits to *C. butyricum*. Among these proteins, there is an annotated enterotoxin (OA81_00270), which has the highest homology (99.6% identity) to an SH3/3D protein (WP_024041145.1) in a *C. butyricum* strain isolated from human gut (11). Moreover, two annotated proteins (OA81_15790 and OA81_07195) have sequence similarity to phage holins (WP_003370185.1 and WP_024040534.1) in *C. botulinum*. A previous report has shown that holin-like *tcdE* is required for exporting enterotoxins *tcdA/B* in *C. difficile* (12). Whether the enterotoxin or holin plays a pathogenic role in *C. diffiicle* infection remains to be examined. In conclusion, genome sequencing is a more accurate method for the identification of *C. butyricum*.

Nucleotide sequence accession numbers. This Whole Genome Shotgun project has been deposited in GenBank under the accession no. JSEG00000000. The version described in this paper is the first version, JSEG01000000.

ACKNOWLEDGMENT

This work was supported by the Research Fund for the Control of Infectious Diseases (Commissioned Project Reference Number CU-12-05-01) from the Food and Health Bureau of the Hong Kong Special Administrative Region Government.

REFERENCES

1. **Xin B, Tao F, Wang Y, Gao C, Ma C, Xu P.** 2013. Genome sequence of *Clostridium butyricum* strain DSM 10702, a promising producer of biofuels and biochemicals. Genome Announc. **1**(4):e00563–13. http://dx.doi.org/10.1128/genomeA.00563-13.
2. **Meng X, Karasawa T, Zou K, Kuang X, Wang X, Lu C, Wang C, Yamakawa K, Nakamura S.** 1997. Characterization of a neurotoxigenic *Clostridium butyricum* strain isolated from the food implicated in an outbreak of food-borne type E botulism. J. Clin. Microbiol. **35**:2160–2162.
3. **Seki H, Shiohara M, Matsumura T, Miyagawa N, Tanaka M, Komiyama A, Kurata S.** 2003. Prevention of antibiotic-associated diarrhea in children by *Clostridium butyricum* MIYAIRI. Pediatr. Int. **45**:86–90. http://dx.doi.org/10.1046/j.1442-200X.2003.01671.x.
4. **Hassan KA, Elbourne LDH, Tetu SG, Johnson EA, Paulsen IT.** 2014. Genome sequence of the neurotoxigenic *Clostridium butyricum* strain 5521. Genome Announc. **2**(3):e00632–14. http://dx.doi.org/10.1128/genomeA.00632-14.
5. **Asha NJ, Wilcox MH.** 2002. Laboratory diagnosis of *Clostridium perfringens* antibiotic-associated diarrhoea. J. Med. Microbiol. **51**:891–894.
6. **Polage CR, Solnick JV, Cohen SH.** 2012. Nosocomial diarrhea: evaluation and treatment of causes other than *Clostridium difficile*. Clin. Infect. Dis. **55**:982–989. http://dx.doi.org/10.1093/cid/cis551.
7. **Chevreux B, Wetter T, Suhai S.** 1999. Genome sequence assembly using trace signals and additional sequence information, p 45–56. *In* Giegerich

R, Hofestädt R, Nattkemper TW. (ed), Proceedings of the German Conferenceon Bioinformatics. GBF-Braunschweig and University of Bielefeld, Bielefeld, Germany.

8. **Bonfield JK, Whitwham A.** 2010. Gap5—editing the billion fragment sequence assembly. Bioinformatics **26:**1699–1703. http://dx.doi.org/10.1093/bioinformatics/btq268.
9. **Teng JLL, Ho TCC, Yeung RSY, Wong AYP, Wang H, Chen C, Fung KSC, Lau SKP, Woo PCY.** 2014. Evaluation of 16SpathDB 2.0, an automated 16S rRNA gene sequence database, using 689 complete bacterial genomes. Diagn. Microbiol. Infect. Dis. **78:**105–115. http://dx.doi.org/10.1016/j.diagmicrobio.2013.10.019.
10. **Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O.** 2008. Toward an online repository of standard operating procedures (SOPs) for (meta)genomic annotation. Omics **12:**137–141. http://dx.doi.org/10.1089/omi.2008.0017.
11. **Brown CT, Sharon I, Thomas BC, Castelle CJ, Morowitz MJ, Banfield JF.** 2013. Genome resolved analysis of a premature infant gut microbial community reveals a *Varibaculum cambriense* genome and a shift towards fermentation-based metabolism during the third week of life. Microbiome. **1:**30. http://dx.doi.org/10.1186/2049-2618-1-30.
12. **Govind R, Dupuy B.** 2012. Secretion of *Clostridium difficile* toxins A and B requires the holin-like protein TcdE. PLoS Pathog. **8:**e1002727. http://dx.doi.org/10.1371/journal.ppat.1002727.