Complete Genome Sequence of *Clostridium botulinum* CJ0611A1, a Type A(B) Isolate Associated with an International Outbreak of Botulism from Commercial Carrot Juice

Richard Harris,[a] Forest Dussault,[a] Annika Flint,[a] John W. Austin,[a] Kelly Weedmark[a]

[a]Bureau of Microbial Hazards, Health Canada, Ottawa, Ontario, Canada

Richard Harris, Forest Dussault, Annika Flint, and Kelly Weedmark contributed equally to this work. Author order was determined at random.

ABSTRACT We report the complete genome (3.9-Mb chromosome, 5.9-kb plasmid) of *Clostridium botulinum* CJ0611A1, a type A(B) strain isolated from carrot juice distributed in Canada and linked to an international 2006 foodborne botulism outbreak. This strain encodes a full-length *bont/A1* gene and a truncated *bont/B* gene.

*C*lostridium botulinum is a Gram-positive, spore-forming, anaerobic bacterium capable of producing botulinum neurotoxin (BoNT), the causative agent of botulism. BoNTs fall into one of seven distinct serological types (A to G), each with several subtypes (1). In 2006, an international outbreak of botulism in Canada and the United States was traced to the consumption of commercial pasteurized carrot juice that contained BoNT type A (2, 3). Brand X carrot juice, associated with outbreaks in Georgia and Florida, and brand Y carrot juice, associated with an outbreak in Ontario, were produced by the same manufacturer in the same plant. *C. botulinum* isolated from carrot juice in Georgia (CDC51303) was found to carry the *bont/A1* gene, while a Florida isolate (CDC51348) contained a *bont/A* and an unexpressed *bont/B* gene (4–6). *C. botulinum* CJ0611A1 is a type A strain isolated from brand Y carrot juice recovered from a refrigerator shared by the two Ontario patients (3).

A single-colony pick from the original food specimen was grown on MT-EYE agar (1.5% McClung-Toabe agar [Difco, Tucker, GA], 5% egg yolk extract, and 5% yeast extract [Difco]) and then frozen at −80°C in multiple Microbank (Pro-Lab Diagnostics, Inc.) cryovials without further subculture. A bead from frozen stock was struck onto MT-EYE agar and incubated overnight at 30°C under anaerobic conditions. Cells from a single-colony TPGY (5% tryptone, 0.5% peptone, 0.4% glucose, 2% yeast extract, and 0.1% sodium thioglycolate) inoculate were collected after 24 h of anaerobic growth at 35°C in TPGY, and DNA was extracted using the Zymo Quick-DNA HMW MagBead kit following resuspension in Zymo DNA/RNA Shield. Eluted DNA was treated with RNase A (2 μl at 10 mg/ml, 30 min at 37°C) and purified using >1,500 bp solid-phase reversible immobilization (SPRI) selection (7). Illumina paired-end (2 × 300 bp) Nextera XT library sequencing was performed on a MiSeq instrument (v3 chemistry) according the manufacturer's instructions (Illumina). Raw Illumina reads (863,750 reads, 301 bp average) were processed using Fastp (v 0.20.0) (8) to remove adapter and barcode sequences, correct mismatched bases in overlaps, and filter low-quality reads. MinION sequencing was performed using a PCR barcoding kit (SQK-PBK004) and SQK-LSK108 sequencing kit on a FLO-106 flow cell (Oxford Nanopore). Following Guppy base calling (Guppy GPU v 3.3.3+fa743ab), raw MinION reads (338,142 reads, 6,431 bp average) were quality control (QC) checked (Nanoplot v 1.29.0) (9), and reads of <10 kb were filtered out (Filtlong v 0.2.0; github.com/rrwick/Filtlong). The hybrid assembly (Unicycler v 0.4, normal mode) (10) was generated using quality-filtered Illumina

(8,416,032 reads, 218 bp average) and MinION (74,388 reads, 16,934 bp average) reads, annotated with PGAP v 2020-07-09.build4716 (best-placed reference protein set, GeneMarkS-2) (github.com/ncbi/pgap), and analyzed with QUAST v 5.0.0 (github.com/ablab/quast). Clustal W v2.1 (11) was used to align *bont* genes and cassettes using references from reference 4.

The closed genome (321-fold coverage) has 28.28% GC content and comprises a circular chromosome (3,918,289 bp) and a plasmid (5,926 bp). A total of 3,450 complete protein-coding DNA sequences (CDSs), 93 pseudogenes, and 80 tRNAs are predicted. The CJ0611A1 genome contains a full-length chromosomal *bont/A1* gene (100% nucleotide identity with ATCC 3502 *bont/A1*, GenBank accession number AM412317) within an *ha⁻ orf⁺* cassette, as well as a truncated *bont/B* sequence ~49 kb downstream in an *ha⁺ orf⁻* cassette. Based on these genomic data as well as toxin assays (4), CJ0611A1, like the Florida isolate (CDC51348), contains an unexpressed *bont/B* gene and is therefore a type A(B) strain (3).

Data availability. The CJ0611A1 genome sequence is available in NCBI (GenBank accession numbers CP059677.1 and CP059678.1); the reads have been deposited in the SRA (SRX8894092 and SRX8894093).

ACKNOWLEDGMENTS

REFERENCES

1. Peck MW, Smith TJ, Anniballi F, Austin JW, Bano L, Bradshaw M, Cuervo P, Cheng LW, Derman Y, Dorner BG, Fisher A, Hill KK, Kalb SR, Korkeala H, Lindström M, Lista F, Lúquez C, Mazuet C, Pirazzini M, Popoff MR, Rossetto O, Rummel A, Sesardic D, Singh BR, Stringer SC. 2017. Historical perspectives and guidelines for botulinum neurotoxin subtype nomenclature. Toxins (Basel) 9:38. https://doi.org/10.3390/toxins9010038.

2. Centers for Disease Control and Prevention (CDC). 2006. Botulism associated with commercial carrot juice: Georgia and Florida. MMWR Morb Mortal Wkly Rep 55:1098–1099.

3. Sheth AN, Wiersma P, Atrubin D, Dubey V, Zink D, Skinner G, Doerr F, Juliao P, Gonzalez G, Burnett C, Drenzek C, Shuler C, Austin J, Ellis A, Maslanka S, Sobel J. 2008. International outbreak of severe botulism with prolonged toxemia caused by commercial carrot juice. Clin Infect Dis 47:1245–1251. https://doi.org/10.1086/592574.

4. Raphael BH, Luquez C, McCroskey LM, Joseph LA, Jacobson MJ, Johnson EA, Maslanka SE, Andreadis JD. 2008. Genetic homogeneity of *Clostridium botulinum* type A1 strains with unique toxin gene clusters. Appl Environ Microbiol 74:4390–4397. https://doi.org/10.1128/AEM.00260-08.

5. Reddy NR, Marshall KM, Morrissey TR, Loeza V, Patazca E, Skinner GE, Krishnamurthy K, Larkin JW. 2013. Combined high pressure and thermal processing on inactivation of type A and proteolytic type B spores of *Clostridium botulinum*. J Food Prot 76:1384–1392. https://doi.org/10.4315/0362-028XJFP-12-538.

6. Marshall KM, Nowaczyk L, II, Raphael BH, Skinner GE, Rukma Reddy N. 2014. Identification and genetic characterization of *Clostridium botulinum* serotype A strains from commercially pasteurized carrot juice. Food Microbiol 44:149–155. https://doi.org/10.1016/j.fm.2014.05.009.

7. Hosomichi K, Mitsunaga S, Nagasaki H, Inoue I. 2014. A bead-based normalization for uniform sequencing depth (BeNUS) protocol for multi-samples sequencing exemplified by HLA-B. BMC Genomics 15:645–2164. https://doi.org/10.1186/1471-2164-15-645.

8. Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

9. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34:2666–2669. https://doi.org/10.1093/bioinformatics/bty149.

10. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.

11. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948. https://doi.org/10.1093/bioinformatics/btm404.