JB Journals.ASM.org

GENOME ANNOUNCEMENT

Draft Genome Sequence of *Clostridium sporogenes* PA 3679, the Common Nontoxigenic Surrogate for Proteolytic *Clostridium botulinum*

Mark Bradbury,[a,b] Paul Greenfield,[c] David Midgley,[a] Dongmei Li,[a] Nai Tran-Dinh,[a] Frank Vriesekoop,[b,d] and Janelle L. Brown[a]

CSIRO Division of Food and Nutritional Sciences, North Ryde, New South Wales, Australia[a]; School of Health Sciences, University of Ballarat, Ballarat, Victoria, Australia[b]; CSIRO Division of Mathematics, Informatics and Statistics, Macquarie University, North Ryde, New South Wales, Australia[c]; and Department of Food Science, Harper Adams University College, Newport, England[d]

***Clostridium sporogenes* PA 3679 is widely used as a nontoxigenic surrogate for proteolytic strains of *Clostridium botulinum* in the derivation and validation of thermal processes in food. Here we report the draft assembly and annotation of the *C. sporogenes* PA 3679 genome. Preliminary analysis demonstrates a high degree of relatedness between *C. sporogenes* PA 3679 and sequenced strains of proteolytic *C. botulinum*.**

The strain of *Clostridium sporogenes* designated Putrefactive Anaerobe (PA) 3679 (ATCC 7955, NCTC 8594) was originally isolated from spoiled canned corn in 1927 (work of E. J. Cameron, as cited in reference 8). It was rapidly adopted as the thermal processing surrogate for proteolytic *Clostridium botulinum* (5, 8) because of morphological similarity and nontoxigenicity and because the heat resistance of its spores exceeded that of *C. botulinum* spores (6, 7). Despite the long history of its use as a surrogate, the genetic basis for the nontoxigenicity of PA 3679 is undetermined (2). To generate knowledge in this regard and provide a basis for exploring the relationships that exist between PA 3679 and proteolytic strains of *C. botulinum*, here we present a draft assembly of the genome of PA 3679.

Genomic DNA of the organism was subjected to Illumina paired-end sequencing. A 4,172,769-bp assembly with ~200× coverage was constructed using Velvet 0.7.63 (9). The assembly consists of 250 large contigs (>200 bp), with a mean GC content of 27.8%. Contigs were arranged into 107 scaffolds, with a N50 contig size of ~44 kb and a maximum contig length of 193,726 bp. Annotation performed by the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (http://www.ncbi.nlm.nih.gov/genomes/static/Pipeline.html) predicted a total of 4,053 protein-coding genes and 45 structural RNAs. Assembly also revealed the presence of an induced prophage (38,651 bp) with ~70× greater sequence coverage than the bacterial genome.

Calculation of the percent shared k-mers ($k = 25$ bp) between PA 3679 and completed genomes of *C. botulinum* revealed a >85% match to type A1 botulinum neurotoxin (BoNT)-producing proteolytic *C. botulinum* strains. In contrast, the k-mer match to nonproteolytic *C. botulinum* was ~15%. BLAST (1) results of the 16S rRNA gene sequence indicated 99 to 100% nucleotide similarity between PA 3679 and a number of proteolytic *C. botulinum* strains, as well as other *C. sporogenes* strains. Multilocus sequence typing (MLST) using seven proteolytic *C. botulinum* housekeeping genes (http://pubmlst.org/cbotulinum) revealed that PA 3679 resides in the same clade as the outlier A1 toxin-producing *C. botulinum* strain A207 (3).

No BoNT-encoding gene clusters or remnants of clusters are present in PA 3679. The regions flanking the A1 BoNT cluster in A1 toxin-producing *C. botulinum* were present in PA 3679; however, they circumscribed a region containing genes with a high degree of similarity to those encoding a putative acetyltransferase and an isochorismatase-like hydrolase present in *C. sporogenes* ATCC 15579 (NZ_ABKW00000000).

Preliminary analysis of the draft genome of PA 3679 highlights the incongruous nature of the taxonomy of *C. botulinum* and *C. sporogenes*. We recommend that the separation of these taxa be revisited, as contemporary analysis of their phylogenetic relationship indicates that their current separation at the species level is likely untenable (4). Completion of the PA 3679 genome and subsequent analysis will clarify understanding of the phylogeny of PA 3679, the evolution of toxigenicity within this branch of the clostridia, and the appropriateness of using PA 3679 as a surrogate and research model for proteolytic *C. botulinum*.

Nucleotide sequence accession numbers. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. AGAH00000000. The version described in this paper is the first version, AGAH01000000.

ACKNOWLEDGMENTS

We gratefully acknowledge the assistance of Scott Chandry, Simon Gladman, the Victorian Bioinformatics Consortium, and Micromon.

This study was funded by CSIRO Food and Nutritional Sciences.

REFERENCES

1. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410.
2. **Brown JL, Tran-Dinh N, Chapman B.** PA 3679 and its uses in the derivation of thermal processing schedules for low-acid shelf-stable foods and as a research model for proteolytic *Clostridium botulinum*. J. Food Prot., in press.
3. **Jacobson MJ, Lin G, Whittam TS, Johnson EA.** 2008. Phylogenetic analysis of *Clostridium botulinum* type A by multi-locus sequence typing. Microbiol. **154**:2408–2415.
4. **Kalia VC, et al.** 2011. Analysis of the unexplored features of rrs (16s rDNA) of the genus *Clostridium*. MBC Genomics **12**:18.

Received 19 December 2011 **Accepted** 4 January 2012

Address correspondence to Mark Bradbury, mark.bradbury@csiro.au.

doi:10.1128/JB.06765-11

5. **McClung LS.** 1937. Studies on anaerobic bacteria X. Heat stable and heat labile antigen in the botulinus and related groups of spore-bearing anaerobes. J. Infect. Dis. **60**:122–128.
6. **National Canners Association Research Laboratories.** 1968. Laboratory manual for food canners and processors, 3rd ed, vol. 1. Microbiology and processing. The AVI 24 Publishing Company, Inc., Westport, CT.
7. **Townsend CT, Somers II, Lamb FC, Olson NA.** 1956. A laboratory manual for the canning industry (2nd ed). National Canners Association Research Laboratories, Washington, DC.
8. **Townsend CT, Esty JR, Baselt FC.** 1938. Heat-resistance studies on spores of putrefactive anaerobes in relation to determination of safe processes for canned foods. J. Food. Sci. **3**:323–346.
9. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.