

Genome Sequence of the Group III *Clostridium botulinum* Strain Eklund-C

Karl A. Hassan,^a Sasha G. Tetu,^a Liam D. H. Elbourne,^a Eric A. Johnson,^b Ian T. Paulsen^{a,c}

Department of Chemistry and Biomolecular Sciences, Macquarie University, NSW, Australia^a; Department of Bacteriology, Botulinum Toxins Laboratory, University of Wisconsin, Madison, Wisconsin, USA^b; J. Craig Venter Institute, Rockville, Maryland, USA^c

The neurotoxins produced by *Clostridium botulinum* strains are among the world's most potent toxins and are the causative agents of paralytic botulism. Here, we present the draft genome sequence of the group III *C. botulinum* strain Eklund-C, including a pseudodysogen-like bacteriophage that harbors the type C neurotoxin operon.

Received 20 January 2013 Accepted 5 February 2013 Published 14 March 2013

Citation Hassan KA, Tetu SG, Elbourne LDH, Johnson EA, Paulsen IT. 2013. Genome sequence of the group III *Clostridium botulinum* strain Eklund-C. *Genome Announc.* 1(2): e00044-13. doi:10.1128/genomeA.00044-13.

Copyright © 2013 Hassan 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 Ian T. Paulsen, ian.paulsen@mq.edu.au.

Clostridium botulinum strains are formidable human and animal pathogens that produce botulinum neurotoxin (BoNT), the most potent neurotoxin described to date and the causative agent of botulism. Animal botulism cases are commonly associated with group III *C. botulinum* strains, which harbor type C or D BoNT genes. Here, we report the draft genome sequence of the prototypical *C. botulinum* group III BoNT/C-producing Eklund strain (Eklund-C). Genomic DNA was sequenced using a Sanger shotgun method to 8-fold coverage and was assembled and annotated as described previously (1). The assembled genome sequence contains 2.96 Mb of sequence data containing 2,578 putative protein-coding genes, 251 putative pseudogenes, and 126 structural RNA genes.

Group III *C. botulinum* strains are phylogenetically related to strains *Clostridium novyi* and *Clostridium haemolyticum* (2, 3). Genome sequences for two strains within this phylogenetic group have been published, those of *C. botulinum* C/D 08-BKT015925 (4) and *C. novyi* NT (5, 6). Approximately 75 to 78% of the annotated protein-coding genes or pseudogenes within the Eklund-C strain were conserved in the 08-BKT015925 and NT strains, that is, 2,206 and 2,132 genes or pseudogenes, respectively, as determined using tBLASTx 2.2.25+ (7) searches (E value $\leq 1e - 30$). Furthermore, the majority of the Eklund-C scaffold sequences could be aligned to the 08-BKT015925 and NT chromosomal sequences using Mauve 2.3.1 (8), highlighting the synteny among the three genomes.

A primary differentiating factor between the toxigenic group III *C. botulinum* strains and closely related species, such as *C. novyi*, is their carriage of BoNT/C or BoNT/D genes on an episomal pseudodysogen-like prophage (9, 10). Five scaffolds containing a total of 177 kb of sequence data are predicted to comprise the Eklund-C BoNT/C bacteriophage genome, which can be aligned to the BoNT/C-encoding phage isolated from strains *C. botulinum* 08-BKT015925 (4) and *C. botulinum* C-Stockholm (11). The sequenced Eklund-C BoNT/C phage genome displays features of a circular self-replicating episomal element: (i) the scaffolds bear no similarity to the *C. novyi* NT chromosomal sequence, (ii) it contains only one copy of a putative terminal direct

repeat sequence that is spanned by numerous sequencing reads and is predicted to mediate circularization of the lysogen, as with the c-st phage (11), and (iii) the two halves of the Eklund-C phage sequence display distinctly different gene orientations and GC skews that diverge abruptly at the putative replication origin and terminus.

The putative Eklund-C phage contains 198 predicted protein-coding genes or pseudogenes. Of these sequences, 135 and 99 were identified within the 08-BKT015925 and C-Stockholm phage genomes, respectively. Most conserved genes are found in syntenous blocks of sequence that are interspersed by novel regions encoding primarily hypothetical proteins. The most highly conserved phage genes include those within the BoNT operon (CBC_A0165-A0170), as well as several phage structural genes (12). BoNT/C- or BoNT/D-encoding phages typically carry a large number of insertion sequence (IS) elements (6). At least 12 intact or degraded IS elements were identified in the Eklund-C prophage, all of which belong to the IS200/605 family. Most of these are related to the elements observed in the 08-BKT015925 and C-Stockholm phages; however, the locations and copy numbers of many of these elements are varied.

Nucleotide sequence accession numbers. The draft genome sequence for *C. botulinum* Eklund-C has been included in the GenBank Whole-Genome Shotgun (WGS) database under the accession no. [ABDQ000000000](http://www.ncbi.nlm.nih.gov/GenBank/ABDQ000000000). The version described is accession no. [ABDQ000000000.1](http://www.ncbi.nlm.nih.gov/GenBank/ABDQ000000000).

ACKNOWLEDGMENTS

C. botulinum Eklund-C was obtained from Mel W. Eklund.

This work was supported by the National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health, Department of Health and Human Services, under NIAID contract no. N01-AI-30071, and by a Life Science Award to I.T.P. from the Office of Science and Medical Research, NSW government, Australia.

REFERENCES

1. Fraser CM, Casjens S, Huang WM, Sutton GG, Clayton R, Lathigra R, White O, Ketchum KA, Dodson R, Hickey EK, Gwinn M, Dougherty B, Tomb JF, Fleischmann RD, Richardson D, Peterson J, Kerlavage AR,

- Quackenbush J, Salzberg S, Hanson M, van Vugt R, Palmer N, Adams MD, Gocayne J, Weidman J, Utterback T, Watthey L, McDonald L, Artiach P, Bowman C, Garland S, Fuji C, Cotton MD, Horst K, Roberts K, Hatch B, Smith HO, Venter JC. 1997. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. *Nature* 390:580–586.
2. Hill KK, Smith TJ, Helma CH, Ticknor LO, Foley BT, Svensson RT, Brown JL, Johnson EA, Smith LA, Okinaka RT, Jackson PJ, Marks JD. 2007. Genetic diversity among botulinum neurotoxin-producing clostridial strains. *J. Bacteriol.* 189:818–832.
 3. Peck MW. 2009. Biology and genomic analysis of *Clostridium botulinum*. *Adv. Microb. Physiol.* 55:183–265, 320.
 4. Skarin H, Hafstrom T, Westerberg J, Segerman B. 2011. *Clostridium botulinum* group III: a group with dual identity shaped by plasmids, phages and mobile elements. *BMC Genomics* 12:185.
 5. Dang LH, Bettgowda C, Huso DL, Kinzler KW, Vogelstein B. 2001. Combination bacteriolytic therapy for the treatment of experimental tumors. *Proc. Natl. Acad. Sci. U. S. A.* 98:15155–15160.
 6. Bettgowda C, Huang X, Lin J, Cheong I, Kohli M, Szabo SA, Zhang X, Diaz LA, Jr, Velculescu VE, Parmigiani G, Kinzler KW, Vogelstein B, Zhou S. 2006. The genome and transcriptomes of the anti-tumor agent *Clostridium novyi*-NT. *Nat. Biotechnol.* 24:1573–1580.
 7. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402.
 8. Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14:1394–1403.
 9. Eklund MW, Poysky FT, Meyers JA, Pelroy GA. 1974. Interspecies conversion of *Clostridium botulinum* type C to *Clostridium novyi* type A by bacteriophage. *Science* 186:456–458.
 10. Ło M, Wegrzyn G. 2012. Pseudolysogeny. *Adv. Virus Res.* 82:339–349.
 11. Sakaguchi Y, Hayashi T, Kurokawa K, Nakayama K, Oshima K, Fujinaga Y, Ohnishi M, Ohtsubo E, Hattori M, Oguma K. 2005. The genome sequence of *Clostridium botulinum* type C neurotoxin-converting phage and the molecular mechanisms of unstable lysogeny. *Proc. Natl. Acad. Sci. U. S. A.* 102:17472–17477.
 12. Hwang HJ, Lee JC, Yamamoto Y, Sarker MR, Tsuchiya T, Oguma K. 2007. Identification of structural genes for *Clostridium botulinum* type C neurotoxin-converting phage particles. *FEMS Microbiol. Lett.* 270:82–89.