Genetic diversity within the botulinumneurotoxin-producing bacteria and their neurotoxins

K. K. Hill^{1*}, G. Xie¹, B. T. Foley² and T. J. Smith³

¹Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM 87545

²Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545

³Molecular and Translational Sciences Division, United States Army Medical Research Institute of Infectious Diseases, Fort Detrick, MD 21702

*Karen K. Hill, Bioscience Division, MS: M88, Los Alamos National Laboratory, Los Alamos, NM 87545, khill@lanl.gov, ph: (505)667-1309, fax: (505)665-3024

Key words: genetic diversity, botulinum neurotoxin, *Clostridiumbotulinum*

Abstract

The recent availability of multiple *Clostridium botulinum* genomic sequences has initiated a new genomics era that strengthens our understanding of the bacterial species that produce botulinum neurotoxins (BoNTs)). Analysis of the genomes has reinforced the historical Group I-VI designations and provided evidence that the *bont* genes can be located within the chromosome, phage or plasmids. The sequences provide the opportunity to examine closely the variation among the toxin genes, the composition and organization of the toxin complex, the regions flanking the toxin complex and the location of the toxin within different bacterial strains. These comparisons provide evidence of horizontal gene transfer and site-specific insertion and recombination events that have contributed to the variation observed among the neurotoxins. Here, examples that have contributed to the variation observed in serotypes A-H strains are presented to illustrate the mechanisms that have contributed to their variation.

Introduction

Clostridium botulinum is a species that is defined by the production of any one of eight serologically distinct botulinum neurotoxins designated BoNT/A-H (Peck, 2009). The toxin is a zinc metalloprotease that consists of a light chain, heavy chain and translocation domain. The three domains interact with accessory proteins of the toxin complex to assist the holotoxin in entering the bloodstream, then neuronal cells (Fujinaga et al., 2013). Once inside the neuronal cells the light chain domain cleaves specific sites within SNARE proteins. The cleavage of these proteins prevents the release of acetylcholine into the neuronal muscular junction and results in a flaccid paralysis known as botulism (Schiavo et al., 1995).

The BoNT-producing bacteria are generally called *C. botulinum*, but they could be considered six different species by 16S *rrn* comparisons (Hill et al., 2009). Group I-IV designations are used to distinguish the different species within *C. botulinum*. The *C. botulinum* Group I strains produce BoNT/A, /B, /F or /H; *C. botulinum* Group II strains produce BoNT/B, /E or /F; *C. botulinum* Group III strains produce BoNT/C or /D; *C. botulinum* Group IV (also designated *C. argentinense*) strains produce BoNT/G. Group V retained the species name of *C. baratii* for strains that produce BoNT/F and Group IV are *C. butyricum* strains that produce BoNT/E (Popoff and Marvaud, 1999). The presence of the same toxin type in different Groups/species indicates that the toxin (and its associated genes within the toxin complex) move into different bacterial species by horizontal gene transfer (HGT).

The eight BoNTs (A-H) proteins are serologically distinct and differ by 37 to 70% in amino acid sequence (Dover et al., 2014b; Hill and Smith, 2013). BoNT/A, /B, /E and /F are the toxins that most commonly cause human botulism. Strains can produce a single toxin or multiple toxins, such as Ba, Bf, Bh, Ab, Af, where the capital letter designates the toxin produced in greater amounts. Six of the eight toxin types (BoNT/A, /B, /C, /D, /E and /F) have variants called subtypes. There are currently 40 recognized toxin subtypes (Figure 1); amino acid

differences may be minor (1-7% difference among type B) or major (3-36% among type F)(Hill and Smith, 2013). The subtypes are given the toxin letter designation followed by a number. For example, there are eight different subtypes of BoNT/A, designated BoNT/A1-A8. Understanding the variation within a toxin type is important to ensure the effectiveness of detection methods or the effectiveness of toxin treatments.

Figure 1. Dendrogram of the BoNT and toxin subtypes. The protein sequences of BoNT/A-H and subtypes illustrate the variation between toxin types and subtypes.

The BoNT protein naturally associates with a several nontoxic proteins to form a toxin complex. Within the genome, the genes for these proteins form one of two different toxin gene clusters: 1) the toxin gene cluster containing a nontoxin-nonhemagglutinin (*ntnh*) gene, a regulatory gene (*botR*) (Connan et al., 2013), a gene of unknown function (*p47*) (Chen et al., 2008) and three predicted open reading frames - this is known as the *orfX*+ gene cluster; and 2) the hemagglutinin (*ha*+) toxin gene cluster, where three hemagglutinin genes replace the *orfX1-X3* genes (Figure 2)(Hill and Smith, 2013). Each of the toxin gene cluster types is associated with specific BoNTs. The *orfX*+ gene cluster is associated with four toxin genes: *bont/A*, */E*, */F* or */H* and the *ha*+ gene cluster is associated with five toxin genes: *bont/A*, */B*, */C*, */D* or */G*. Only the *bont/A* is found within either the *orfX*+ or *ha*+ gene cluster. All other *bont* genes are located within one or the other of the toxin gene clusters.

Figure 2. The two different toxin gene clusters within botulinum neurotoxin-producing strains. The toxin gene clusters within the *bont*-containing clostridia code for different proteins that surround and protect the botulinum neurotoxin. The upper part of the figure shows the *orfX*+ toxin gene cluster. The lower portion of the figure shows three different arrangements of *ha*+ gene clusters.

The toxin type and subtype, the presence of its gene within an *aha+* or *orfX+* gene cluster, and its location within various bacterial species/Groups provide valuable information related to characterization of the toxin and the bacteria that producing these toxins. The same toxin gene's presence in different bacterial species illustrates insertion or HGT events. The changes in toxin sequences through recombination events can result in the creation of new subtypes. In addition, the genomic location of the toxin genes does not appear to be random and examination of the mechanisms of toxin gene insertions can assist in understanding relationships of toxin-containing strains.

HGT of the toxin gene can be facilitated by different mechanisms, including the presence of the toxin within mobile elements such as plasmids or phage, the association of the toxin with transposases or Insertion Sequence (IS) elements, or the presence of direct repeated sequences in the regions flanking the toxin gene cluster which may provide a targeting and insertion site. The mobility of the toxin gene cluster and its presence in different bacterial species provides the opportunity for the toxins to evolve within different bacterial backgrounds and environments through genetic recombination, resulting in the diverse array of toxins that have been identified. Below are descriptions of significant insertion and recombination events that have resulted in toxin diversity or the placement of toxin genes within different bacterial Groups/species. The examples presented here are not comprehensive descriptions of each toxin but rather provide examples illustrating various toxin mobility, insertion, and recombination mechanisms in multiple species of bacteria that result in botulinum neurotoxin diversity.

Results

BoNT/A

BoNT/A is produced by *C. botulinum* Group I bacteria and currently consists of 8 subtypes (BoNT/A1-A8) that have protein differences that are greater than 15% (Kull et al., 2015; Smith, 2014). The location of the *bont/A* gene within a strain can be within one of two sites within the

chromosome (*arsC* operon or *oppB/brnQ* operon), or at one site within large Group I plasmids (Hill et al., 2009). Comparisons of the genomic sequences in multiple strains have shown that the sites of the *bont* genes within a plasmid or chromosome are not random. Figure 3 shows that the *bont/A* subtype genes are located within the *arsC* operon. For example, several *orfX*⁺ gene clusters that contain *bont/A1*, */A2*, or */A8* are located among genes that form part of the *arsC* operon (Hill et al., 2009; Kull et al., 2015). The *oppB/brnQ* operon contains *bont/A1* or */A5* genes that form part of the *ha*⁺ gene cluster (Carter et al., 2011; Hill et al., 2009).

Locations of the *bont* genes within the Group I plasmids are also conserved (Figure 3). The Group I plasmids vary in size (~150-270 kb), but often show significant regions of synteny. In addition to the *bont/A* genes located within Group I plasmids, there is a second location that is sometimes associated with *bont/B* toxin genes. The *bont/A* site within Group I plasmids can contain *bont/A2*, */A3* or */A4*, as well as occasional *bont/F* gene clusters. The *bont/B* site has been associated with *bont/B1*, */B2*, and */B5* genes. The recent genomic sequences of bivalent BoNT/A2b5 strain CDC 1436 and BoNT/A2b3 strain 87 (Giordani et al., 2014; Smith et al., 2015) show the site of their *bont/A2* genes within plasmids, in contrast to the *arsC* operon chromosomal location of the *bont/A2* gene from the Kyoto-F strain. Comparisons of the *bont/A2b5*- and */A2b3*-containing plasmids show significant shared synteny with each other and to the plasmid in the bivalent BoNT/B5a4-producing *C. botulinum* strain 657.

123

124

Group I

Group II

Group IV

Group V

Figure 3. Relative location of toxin gene clusters within chromosome or plasmids. The location of the different *bont* genes and their toxin cluster type (*ha+* or *orfX+*) are shown in Group I, II, IV and V strains.

Interestingly, the *bont/A1* gene is the only toxin subtype gene that is located within either *ha+* or *orfX+* toxin gene clusters. This is the result of a recombination event within the nontoxic nonhemagglutinin (*ntnh*) gene that precedes the *bont* gene in all toxin gene clusters. The *ntnh* gene within serotype B strains (within a *ha+* toxin gene cluster) recombined with a serotype A *ntnh* gene, resulting in a B/A *ntnh* chimera (East et al., 1996; Hill et al., 2009). The recombination event resulted in the placement of the *bont/A1* gene after the B/A *ntnh* chimera, within the *ha+* gene cluster. This B/A *ntnh* recombinant event resulted in a very successful lineage of HA+ BoNT/A1 strains. The *bont/A5* subtype is also found within the *ha+* toxin gene cluster, indicating that BoNT/A5 strains are closely related to HA+ BoNT/A1 strains. The *orfX+* *bont/A1* gene cluster contains the ancestral *ntnh* gene associated with *orfX+* BoNT/A1 strains, as well as strains that produce BoNT/A2, /A3, /A4, and /A8 toxins (Hill et al., 2009; Kull et al., 2015). Another significant recombination event in serotype A strains involved the recombination between two *bont* genes to create a new subtype. The *bont/A2* is the result of a recombination event between the *bont/A1* and *bont/A3* genes (Hill et al., 2007).

BoNT/B

In contrast to BoNT/A, which is produced only within Group I strains, BoNT/B is produced by both *C. botulinum* Group I and Group II strains. The proteins of the eight BoNT/B subtypes (B1-B8) differ by 7% (Smith, 2014; Wangroongsarb et al., 2014). As with the serotype A strains, the *bont/B* gene can be located within a plasmid or chromosome. For example, the majority of *bont/B1* genes are found within plasmids, but most *bont/B2* genes are located within the chromosome (Franciosa et al., 2009). The chromosomal location of the *bont/B2* within the

Bh strain, IBCA10-0760, is at the *oppA/brnQ* site (Figure 3)(Dover et al., 2014b). This is also the location of the unexpressed silent B (designated (B)) gene and the *ha+ bont/A1*.

Bont/B genes have also been located within the Group I *bont*-containing plasmids such as *bont/B1* (Okra), */B2* (Prevot 594) and */B5* (bivalent B strains). The plasmid-borne *bont/B1*, */B2* and */B5* share a similar site when the Group I plasmids are compared. The plasmid location of the *bont/B2* gene in the recently sequenced Prevot 594 strain provides another example where the same toxin subtype can be within either a chromosomal or plasmid location (Smith et al., 2015).

In *C. botulinum* Group II strains, the *bont/B4* gene may be located within one of three ~47-63 kb plasmids that do not share synteny with the Group I *bont*-containing plasmids (Carter et al., 2014). All BoNT/B4 strains that have been examined to date appear to carry the *bont/B4* genes within these plasmids.

BoNT/C and BoNT/D

BoNT/C and BoNT/D are produced by *C. botulinum* Group III strains and their *bont* genes are located within phage. The toxin genes of *bont/C* and *bont/D* have recombined to create mosaics or chimeric BoNTs that are designated as BoNTC/D and BoNTD/C. The chimeric BoNTs of C/D and D/C are considered variants or subtypes, that differ by up to 23% at the amino acid level. Examination of the chimeric D/C or C/D sequences shows that the mosaics are composed of approximately 2/3 one serotype and 1/3 of the other serotype (Webb et al., 2007). Serotypes C and D only very rarely affect humans and are predominantly associated with botulism in birds and other waterfowl (C/D) or various herbivores, such as cattle or horses (C, D, D/C).

While *bont/C1* toxin genes are within *ha+* gene clusters, the gene cluster arrangement differs from others in that the *botR* transcriptional regulator gene has moved to the 5' end of the cluster and the orientation is now 5' to 3', instead of the 3' to 5' direction observed in other *ha+* gene clusters (Figure 2)(Hauser et al., 1994; Sakaguchi et al., 2005). This illustrates that

recombination or HGT events may occur among the nontoxic genes of the toxin cluster as well as within the *bont* genes.

BoNT/E

BoNT/E is produced by *C. botulinum* Group II or *C. butyricum* Group VI type E strains and consists of 12 subtypes. The BoNT/E1-E12 proteins vary by up to 11% (Mazuet et al., 2015; Smith, 2014; Weedmark et al., 2014). Serotype E strains are almost always isolated in northern geographic regions, such as Alaska/Canada, northern Japan, and the Baltic Sea region. A notable exception was the BoNT/E9-producing strain CDC 66177 (Raphael et al., 2012), which was isolated from the environment in southern Argentina. More recently, BoNT/E10 and BoNT/E11 strains were isolated from the eastern region of Hudson Bay and Ungava Bay (E10) or in the vicinity of the Koksoak River (E11) in Canada (Weedmark et al., 2014). The BoNT/E12 strain was recently identified from a foodborne botulism case in France (Mazuet et al., 2015).

Genetic analysis of the strains producing BoNT/E9 and /E12 indicates that their genomes are more closely related to the Group II BoNT/B4 Eklund 17B strains than to other serotype E strains (Mazuet et al., 2015; Raphael et al., 2012).

The *bont/E* gene can be located within the chromosome or within an ~146kb plasmid (Zhang et al., 2013). The few genomic sequences of the *C. butyricum* Group VI show a chromosomal location for the *bont/E*. The presence of the *bont/E* gene in both *C. botulinum* Group II and *C. butyricum* Group VI strains results from an insertion event involving a resolvase, encoded by the *rarA* gene that is a recombinase associated with transposases (Figure 3). In the BoNT/E-producing *C. botulinum* and *C. butyricum*, the *rarA* gene appears to find a *rarA* homolog. The *rarA* homolog is split and an intact copy of the *rarA* plus the nearby *bont/E* toxin cluster is inserted. The result is an inserted DNA segment that contains an intact *rarA* gene and the *bont/E* toxin cluster bordered by a 5' partial *rarA* gene and a 3' partial *rarA* gene. The intact *rarA* gene within the inserted DNA, to include the intact *rarA* gene, appears to be very

similar in both *C. butyricum* and *C. botulinum* strains, indicating a common source(Hill et al., 2009). However splicing together of the 5' and 3' split *rarA* homolog sequences show that they are different in the two species (*C. butyricum* and *C. botulinum*) indicating separate insertion events (Hill et al., 2009).

BoNT/F

BoNT/F is produced by three species of bacteria, Group I or Group II *C. botulinum*, or *C. baratii* Group V strains. The protein sequence variation between the BoNT/F1-F7 subtypes may be up to 36%(Smith, 2014),making these the most diverse subtypes within any serotype.In Group I strains, the *bont/F1* gene shares the same *arsC* operon chromosome location as is observed with the *orfX+bont/A1* or */A2*.However, the *bont/F4* is located in the *pulE* site of the chromosome, while the *bont/F2* and */F5* are located within a plasmid (in the same site as the *bont/A2*, */A3* or */A4* in Group I plasmids) (Figure 3). In *C. botulinum* Group II strains, the *bont/F6* gene is located at a chromosomal site designated *topB*, for the gene encoding topoisomerase BIII(Carter et al., 2013). Within the Group V *C. baratii*, the *bont/F7* is also found within the chromosome.

The serotype F strains illustrate different examples of insertion and recombination events. The *bont/F4* and *bont/F6* genes share a similar insertion mechanism observed in serotype E strains, where the *bont/E* is inserted with a recombinase associated with transposases. In Group II *C. botulinum* *bont/F6* strains, it is the *topB* gene that splits a gene homolog, inserting a segment of DNA including the intact *topB* gene and the *bont/F6* gene cluster along with other DNA(Carter et al., 2013). The insertion site is bordered by the 5' split *topB* and 3' split *topB*. Likewise, the *bont/F4* is located at the *pulE* chromosomal site in Group I strains. *PulE* is a gene that codes for an ATPase PilB-like protein that is part of the pilus assembly pathway. Here a *pulE* homolog is split, inserting the *bont/F4* complex between the 5' and 3' split *pulE*(Dover et al., 2013).It is of interest that the same insertion mechanism is used by multiple BoNT/E and /F subtypes from *C. botulinum* Group I and Group II, and Group VI *C. butyricum* strains.

Similar to the A1-A3 recombination event that created the A2 subtype, a recombination event between two *bont/F* subtype genes has created a new subtype. The *bont/F6* is the result of a recombination event between the *bont/F1* and the *bont/F2* (Smith et al., 2015). Importantly, the *bont/F6* gene is found within Group II *C. botulinum* strains, yet the *bont/F1* and */F2* are found within Group I *C. botulinum* strains, indicating not only a recombination event has occurred, but also its presence in Group II *C. botulinum* strains indicates an interspecies HGT event.

The serotype F strains also provide an example where direct repeats of intact IS elements could facilitate the insertion of the *bont/F7* into the *C. baratii* type F strains. Recently the genome of a *bont/F7*-producing *C. baratii* type V strain Sullivan was sequenced. Examination of the chromosomal sequence showed that the *bont/F7* toxin cluster was flanked by direct repeats of IS 1182 elements that share 99% identity to each other (Dover et al., 2014a; Smith et al., 2015). The direct repeats of the intact IS elements may explain the movement of the *bont/F7* gene into the *C. baratii* species.

BoNT/G

BoNT/G is produced by the Group IV *C. botulinum*, currently designated as Group IV *C. argentinense*. BoNT/G-producing strains are rare and only recently has a genome been sequenced, from strain CDC 2741. Analysis of the *bont/G* toxin cluster sequence revealed it is within an *ha+* gene cluster that contains all three hemagglutinin genes (*ha70*, *ha17*, and *ha33*), but the gene arrangement is unlike other *ha+* toxin clusters, with the initial hemagglutinin gene being *ha33*, not *ha70* (Figure 2) (Smith et al., 2015). Examination of the *ha33* gene within the *bont/G* toxin cluster showed it had the greatest similarity to the *ha33* within serotype C or D strains and was more distantly related to *ha 33* within serotype A or B strains. The discovery of the unique arrangement of the *ha33* gene within the *bont/G* toxin gene cluster provides another example of the movement of nontoxic gene components within the toxin gene cluster due to recombination events or HGT.

BoNT/H

BoNT/H is a novel toxin type that was isolated from the bivalent Group I *C. botulinum* strain, IBCA10-7060. It produces both BoNT/B2 and BoNT/H (Dover et al., 2014b). Both of the toxin gene clusters are located within the chromosome. The *bont/B2* gene is located at the *oppB/brnQ* chromosomal site. The *bont/H* gene is found at a unique location within the chromosome resulting from an insertion event facilitated by flanking IS 110 elements that share 99% identity (Dover et al., 2014b). BoNT/H is similar in light chain sequence (79.6% overall amino acid identity) and shares the same enzymatic recognition and cleavage site as BoNT/F5 (Kalb et al., 2015). Its receptor-binding sequence closely resembles BoNT/A (77.3-92.5% identity at the protein level), while the central portion of the toxin molecule, its translocation domain, is only vaguely related to BoNT/F (62.6% identity or less), its closest relative. The partially chimeric nature of this toxin, combined with its sequence divergence, indicates that this may be an ancestral BoNT that has undergone multiple rounds of recombination events. This chimeric nature has also sparked a controversy over its classification. Recent journal articles focusing on the functional characteristics (Kalb et al., 2015) and genetic sequence (Gonzalez-Escalona et al., 2014) of the toxin underscore the controversial nature of its nomenclature.

Conclusion

Examination of the sequences from *bont* genes, toxin complex genes, and genomes from multiple *C. botulinum* strains has enabled us to increase our understanding of the mechanisms underlying toxin diversity. These mechanisms include the movement of genetic material using mobile genetic elements, such as plasmids and phage, and insertion and recombination events facilitated by a variety of insertion mechanisms, including simple recombination events, targeting of specific gene homologs followed by insertion of genetic material, and insertions facilitated by IS elements. For example, a recombination event between serotype B and serotype

Antnh genes removed the *bont/A1* from its *orfX*+ gene cluster and placed it within the *ha*+ toxin cluster, resulting in a highly successful and robust clonal expansion of the HA+BoNT/A1-producing *C. botulinum* strains. Recombination events involving the *bont* genes themselves have created the BoNT/A2 subtype as a chimera of *bont/A1* and /A3, and the BoNT/F6 subtype as a chimera of *bont/F1* and /F2. Recombination of BoNT/C and BoNT/D to produce mosaic C/D and D/C toxins demonstrated that recombination between *bont* genes of different serotypes is possible. BoNT/E and BoNT/F strains show movement of the toxin genes into different Groups/species by their presence near transposase-associated genes. Insertion of *bont/F7* genes in *C. baratii* strains and the *bont/H* gene in IBCA10-7060 strain via direct repeats of IS elements that flank the toxin cluster explain the movement into the different species. These events may be quite recent, due to the intact IS elements with high shared identities. The widespread finding of *bont* toxin clusters within mobile elements, such as plasmids or phage, provides a mechanism for the movement and interaction of this genetic material among various botulinum neurotoxin-producing clostridial species.

The production of novel toxins through recombination with known toxin genes emphasizes the expanding realm of diversity within neurotoxins. New toxin subtypes are constantly being identified and we have now been challenged with a newly identified BoNT/H, indicating that even more genetic variation in the botulinum neurotoxins may remain to be discovered. An understanding of this diversity is necessary to ensure the development of effective diagnostics and countermeasures for intoxication, and may lead to improved therapeutic toxins in the future.

- Carter, A.T., Austin, J.W., Weedmark, K.A., Corbett, C., Peck, M.W., 2014. Three classes of plasmid (47-63 kb) carry the type B neurotoxin gene cluster of group II *Clostridium botulinum*. *Genome biology and evolution* 6, 2076-2087.
- Carter, A.T., Pearson, B.M., Crossman, L.C., Drou, N., Heavens, D., Baker, D., Febrer, M., Caccamo, M., Grant, K.A., Peck, M.W., 2011. Complete genome sequence of the proteolytic *Clostridium botulinum* type A5 (B3') strain H04402 065. *Journal of bacteriology* 193, 2351-2352.
- Carter, A.T., Stringer, S.C., Webb, M.D., Peck, M.W., 2013. The type F6 neurotoxin gene cluster locus of group II *Clostridium botulinum* has evolved by successive disruption of two different ancestral precursors. *Genome biology and evolution* 5, 1032-1037.
- Chen, Y., Korkeala, H., Linden, J., Lindstrom, M., 2008. Quantitative real-time reverse transcription-PCR analysis reveals stable and prolonged neurotoxin cluster gene activity in a *Clostridium botulinum* type E strain at refrigeration temperature. *Applied and environmental microbiology* 74, 6132-6137.
- Connan, C., Deneve, C., Mazuet, C., Popoff, M.R., 2013. Regulation of toxin synthesis in *Clostridium botulinum* and *Clostridium tetani*. *Toxicon : official journal of the International Society on Toxinology* 75, 90-100.
- Dover, N., Barash, J.R., Burke, J.N., Hill, K.K., Detter, J.C., Arnon, S.S., 2014a. Arrangement of the *Clostridium baratii* F7 Toxin Gene Cluster with Identification of a Sigma factor that recognizes the botulinum toxin gene cluster promoters. *PloS one* 9, e97983.
- Dover, N., Barash, J.R., Hill, K.K., Davenport, K.W., Teshima, H., Xie, G., Arnon, S.S., 2013. *Clostridium botulinum* strain Af84 contains three neurotoxin gene clusters: bont/A2, bont/F4 and bont/F5. *PloS one* 8, e61205.
- Dover, N., Barash, J.R., Hill, K.K., Xie, G., Arnon, S.S., 2014b. Molecular characterization of a novel botulinum neurotoxin type h gene. *J Infect Dis* 209, 192-202.
- East, A.K., Bhandari, M., Stacey, J.M., Campbell, K.D., Collins, M.D., 1996. Organization and phylogenetic interrelationships of genes encoding components of the botulinum toxin complex in proteolytic *Clostridium botulinum* types A, B, and F: evidence of chimeric sequences in the gene encoding the nontoxic nonhemagglutinin component. *International journal of systematic bacteriology* 46, 1105-1112.
- Franciosa, G., Maugliani, A., Scalfaro, C., Aureli, P., 2009. Evidence that plasmid-borne botulinum neurotoxin type B genes are widespread among *Clostridium botulinum* serotype B strains. *PloS one* 4, e4829.
- Fujinaga, Y., Sugawara, Y., Matsumura, T., 2013. Uptake of botulinum neurotoxin in the intestine. *Curr Top Microbiol Immunol* 364, 45-59.
- Giordani, F., Anselmo, A., Fillo, S., Palozzi, A.M., Fortunato, A., Gentile, B., Ciammaruconi, A., Anniballi, F., De Medici, D., Lista, F., 2014. GenBank posting of *Clostridium botulinum* A2B3 87 plasmid p1_72A2B3 (AUZB01000012).
- Gonzalez-Escalona, N., Thirunavukkarasu, N., Singh, A., Toro, M., Brown, E.W., Zink, D., Rummel, A., Sharma, S.K., 2014. Draft Genome Sequence of Bivalent *Clostridium botulinum* Strain IBCA10-7060, Encoding Botulinum Neurotoxin B and a New FA Mosaic Type. *Genome announcements* 2.
- Hauser, D., Eklund, M.W., Boquet, P., Popoff, M.R., 1994. Organization of the botulinum neurotoxin C1 gene and its associated non-toxic protein genes in *Clostridium botulinum* C 468. *Molecular & general genetics : MGG* 243, 631-640.
- Hill, K.K., Smith, T.J., 2013. Genetic diversity within *Clostridium botulinum* serotypes, botulinum neurotoxin gene clusters and toxin subtypes, in: Rummel, A., Binz, T. (Eds.), *Botulinum Neurotoxins*. Springer, New York, pp. 1-20.
- Hill, K.K., Smith, T.J., Helma, C.H., Ticknor, L.O., Foley, B.T., Svensson, R.T., Brown, J.L., Johnson, E.A., Smith, L.A., Okinaka, R.T., Jackson, P.J., Marks, J.D., 2007. Genetic diversity among Botulinum Neurotoxin-producing clostridial strains. *Journal of bacteriology* 189, 818-832.
- Hill, K.K., Xie, G., Foley, B.T., Smith, T.J., Munk, A.C., Bruce, D., Smith, L.A., Brettin, T.S., Detter, J.C., 2009. Recombination and insertion events involving the botulinum neurotoxin complex genes in *Clostridium botulinum* types A, B, E and F and *Clostridium butyricum* type E strains. *BMC biology* 7, 66.
- Kalb, S.R., Baudys, J., Raphael, B.H., Dykes, J.K., Luquez, C., Maslanka, S.E., Barr, J.R., 2015. Functional Characterization of Botulinum Neurotoxin Serotype H as a Hybrid of Known Serotypes F and A (BoNT F/A). *Analytical chemistry*.
- Kull, S., Schulz, K.M., Weisemann, J., Kirchner, S., Schreiber, T., Bollenbach, A., Dabrowski, P.W., Nitsche, A., Kalb, S.R., Dorner, M.B., Barr, J.R., Rummel, A., Dorner, B.G., 2015. Isolation and functional characterization of the novel *Clostridium botulinum* neurotoxin A8 subtype. *PloS one* 10, e0116381.
- Mazuet, C., Sautereau, J., Legeay, C., Bouchier, C., Bouvet, P., Popoff, M.R., 2015. An atypical outbreak of food-borne botulism due to *Clostridium botulinum* types B and E from ham. *Journal of clinical microbiology* 53, 722-726.
- Peck, M.W., 2009. Biology and genomic analysis of *Clostridium botulinum*. *Advances in microbial physiology* 55, 183-265, 320.

Popoff, M., Marvaud, J.C., 1999. Structural and genomic features of clostridial neurotoxins, in: Alouf, J.E., Popoff, M.R. (Eds.), *The Comprehensive Sourcebook of Bacterial Protein Toxins*. Academic Press, pp. 174-201.

Raphael, B.H., Lautenschlager, M., Kalb, S.R., de Jong, L.I., Frace, M., Luquez, C., Barr, J.R., Fernandez, R.A., Maslanka, S.E., 2012. Analysis of a unique *Clostridium botulinum* strain from the Southern hemisphere producing a novel type E botulinum neurotoxin subtype. *BMC microbiology* 12, 245.

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. *Proceedings of the National Academy of Sciences of the United States of America* 102, 17472-17477.

Schiavo, G., Rossetto, O., Tonello, F., Montecucco, C., 1995. Intracellular targets and metalloprotease activity of tetanus and botulism neurotoxins. *Curr Top Microbiol Immunol* 195, 257-274.

Smith, T.J., 2014. *Clostridium botulinum* genomes and genetic diversity, in: Foster, K.A. (Ed.), *Molecular Aspects of Botulinum Neurotoxin*. Springer, New York, pp. 207-228.

Smith, T.J., Hill, K.K., Xie, G., Foley, B.T., Williamson, C.H., Foster, J.T., Johnson, S.L., Chertkov, O., Teshima, H., Gibbons, H.S., Johnsky, L.A., Karavis, M.A., Smith, L.A., 2015. Genomic sequences of six botulinum neurotoxin-producing strains representing three clostridial species illustrate the mobility and diversity of botulinum neurotoxin genes. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* 30, 102-113.

Wangroongsarb, P., Kohda, T., Jittaprasartsin, C., Suthivarakom, K., Kamthlang, T., Umeda, K., Sawanpanyalert, P., Kozaki, S., Ikuta, K., 2014. Molecular characterization of *Clostridium botulinum* isolates from foodborne outbreaks in Thailand, 2010. *PloS one* 9, e77792.

Webb, R.P., Smith, T.J., Wright, P.M., Montgomery, V.A., Meagher, M.M., Smith, L.A., 2007. Protection with recombinant *Clostridium botulinum* C1 and D binding domain subunit (Hc) vaccines against C and D neurotoxins. *Vaccine* 25, 4273-4282.

Weedmark, K.A., Lambert, D.L., Mabon, P., Hayden, K.L., Urfano, C.J., Leclair, D., Van Domselaar, G., Austin, J.W., Corbett, C.R., 2014. Two novel toxin variants revealed by whole-genome sequencing of 175 *Clostridium botulinum* type E strains. *Applied and environmental microbiology* 80, 6334-6345.

Zhang, Z., Hintsala, H., Chen, Y., Korkeala, H., Lindstrom, M., 2013. Plasmid-borne type E neurotoxin gene clusters in *Clostridium botulinum* strains. *Applied and environmental microbiology* 79, 3856-3859.

385 **Acknowledgements:** This research was partially funded by a grant from NIH/NIAID U01
386 AI056493. The opinions, interpretations, conclusions, and recommendations are those of
387 the authors and are not necessarily endorsed by the U.S. Army.