Genomic organization and diversity of *Clostridium botulinum* group III

The bug behind animal botulism

Hanna Skarin

Faculty of Veterinary Medicine and Animal Science
Department of Biomedical Sciences and Veterinary Public Health,
Uppsala

and

Department of Bacteriology, National Veterinary Institute, Uppsala

Doctoral Thesis
Swedish University of Agricultural Sciences
Uppsala 2015

Acta Universitatis agriculturae Sueciae

2015:80

Cover: Agar plate with colonies of a *Clostridium botulinum* group III strain (photo: Karl-Erik Johansson, BVF). The images of the DNA spiral and the bacteriophage are obtained from Pixabay.

Genomic organization and diversity of *Clostridium botulinum* group III – the bug behind animal botulism

Abstract

Botulism is caused by botulinum neurotoxins (BoNTs) produced by the spore forming strictly anaerobic bacterium *Clostridium botulinum*. Seven different types of BoNTs (type A-F) have so far been established on the basis of neutralization with different antibodies. Botulism affects both humans and animals, and there are occasionally large-scale outbreaks of high mortality in animals. Especially large outbreaks of avian botulism have been reported from various countries, including Sweden. Other animals relatively commonly affected are cattle, horses, sheep and farmed fur animals. *C. botulinum* is a diverged species and can be organized into four groups, which reflect their genetic and physiological differences. *C. botulinum* group III strains producing BoNT types C, D and chimers C/D and D/C, are mainly connected to animal botulism. The gene encoding BoNT in *C. botulinum* group III strains is located on an unstable plasmid-like phage.

In this thesis, strains of the previously relatively uncharacterised *C. botulinum* group III were isolated and genotyped with pulsed-field gel electrophoresis. Several pulsotypes were formed, but the majority clustered closely together and represented most of the chimeric strains. Strains representing different pulsotypes and different animal and geographical origin, were selected for whole genome sequencing and the resulting genomes could be divided into four genomic lineages. Comparisons against genomes of *Clostridium novyi* and *Clostridium haemolyticum* revealed that they could be organized into the same genomic lineages (lineages II-IV), which resulted in the suggested collective term *C. novyi sensu lato*. The organization of all sequenced genomes was analysed. It revealed a relatively conserved chromosome and an abundance of highly dynamic plasmids. The plasmids, lineages and species were entwined because plasmids and toxin genes had moved across the lineage boundaries. Of the four lineages, only lineage I was *C. botulinum* specific, and this lineage includes strains of the most common pulsotype. One genome of lineage I was assembled into completion. It was smaller than *C. botulinum* group I and II genomes, but contained as much as five plasmids. Most of the identified putative toxin genes were found on these plasmids. Strains of lineage I may be more virulent than other *C. botulinum* group III strains, which is reflected by their domination in animal botulism cases today.

Keywords: Clostridium botulinum, botulinum neurotoxin, BoNT, botulism, PFGE, genome, plasmid, bacteriophage, *Clostridium novyi, Clostridium haemolyticum*.

Author's address: Hanna Skarin, SVA, Department of Bacteriology, 756 51 Uppsala, Sweden. *E-mail:* hanna.skarin@sva.se

Till min morfar som är en riktig biolog

Contents

List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I **Skarin, H**., Lindberg, A., Blomqvist, G., Aspán, A. and Båverud, V. (2010). Molecular characterization and comparison of *Clostridium botulinum* type C avian strains. *Avian Pathology,* 39(6): 511–518.

II Anza, I[1]., **Skarin, H**[1]., Vidal D., Lindberg, A., Båverud, V. and Mateo, R. (2014). The same clade of *Clostridium botulinum* strains is causing avian botulism in southern and northern Europe. *Anaerobe*, 26: 20-23.
[1]These authors contributed equally to this publication

III **Skarin, H**., Håfström, T., Westerberg, J. and Segerman, B. (2011). *Clostridium botulinum* group III: a group with dual identity shaped by plasmids, phages and mobile elements. *BMC Genomics*, 12:185. doi:10.1186/1471-2154-12-185.

IV **Skarin, H**. and Segerman, B. (2014). Plasmidome interchange between *Clostridium botulinum*, *Clostridium novyi* and *Clostridium haemolyticum* converts strains of independent lineages into distinctly different pathogens. *PLoS ONE*, 9 (9): e107777. doi:10.1371/journal.pone.0107777.

Papers I and II are reproduced with the permission of the publisher.

The contribution to the papers included in this thesis was as follows:

I   Hanna Skarin, Anna Lindberg and Viveca Båverud planned and designed the experiments. Hanna Skarin performed the experiments and analysed the results together with Anna Aspán. Hanna Skarin wrote the first draft of the manuscript with input from Anna Lindberg, Gunilla Blomqvist, Viveca Båverud and Anna Aspán.

II Hanna Skarin and Ibone Anza Gomez planned the study. Hanna Skarin performed the Immunomagnetic separation experiments. Ibone Anza Gomez collected samples from Spanish avian botulism outbreaks. Hanna Skarin and Ibone Anza Gomez performed isolation and genotyping experiments and analysed the results. Hanna Skarin and Ibone Anza Gomez wrote the manuscript with input from Dolors Vidal, Rafael Mateo, Anna Lindberg and Viveca Båverud.

III Hanna Skarin and Bo Segerman conceived and designed the experiments. Hanna Skarin, Therese Håfström and Josefina Westerberg performed the experiments. Hanna Skarin, Therese Håfström and Bo Segerman analysed the results. Hanna Skarin and Bo Segerman wrote the manuscript.

IV Hanna Skarin and Bo Segerman conceived and designed the experiments. Hanna Skarin performed the experiments. Hanna Skarin and Bo Segerman analysed the results and wrote the manuscript.

Related publications

Additional publications have been produced during this thesis project, which also relates to the topic:

- Lindberg, A., Skarin, H., Knutsson, R., Blomqvist, G. and Båverud, V. (2010). Real-time PCR for *Clostridium botulinum* type C neurotoxin (BoNTC) gene, also covering a chimeric C/D sequence – application on outbreaks of botulism in poultry. *Veterinary Microbiology*, 146 (1-2), 118-123.
- Skarin, H. and Segerman, B. (2011). Horizontal transfer of toxin genes in *Clostridium botulinum*: Involvement of mobile elements and plasmids. *Mobile Genetic Elements,* 1(3), 213-215.
- Woudstra, C., Skarin, H., Anniballi, F., Fenicia, L., Bano, L., Drigo, I., Koene, M., Bäyon-Auboyer, M.H., Buffereau, J.P., De Medici, D., and Fach, P. (2012). Neurotoxin gene profiling of *Clostridium botulinum* types C and D native to different countries within Europe. *Applied Environmental Micriobiology,* 78(9), 3120-3127.
- Skarin, H., Tevell Åberg, A., Woudstra, C., Hansen, T., Löfström, C., Koene, M., Bano, L., Hedeland, M., Anniballi, F., De Medici, D., Olsson Engvall, E. (2013). The workshop on animal botulism in Europe. *Biosecurity and bioterrorism: biodefense strategy, practice, and science*, 11(Suppl. 1), 183-190.
- Skarin, H., Lindgren, Y., and Jansson, D. (2015). Investigations into an outbreak of botulism caused by *Clostridium botulinum* type C/D in laying hens. *Avian Diseases,* 59, 335-340.

Abbreviations

AFLP   amplified fragment length polymorphism
BoNT   botulinum neurotoxin
bont   botulinum neurotoxin gene
bp   base pair
contig   contiguous sequence
DURC   dual use research of concern
HA   hemagglutinin
HC   heavy chain
LC   light chain
MLST   multilocus sequence typing
MLVA   multiple-locus variable number tandem repeat analysis
NAP   non-toxic neurotoxin-associated proteins,
NGS   next-generation sequencing
NTNH   non-toxic non-hemagglutinin protein
PCR   polymerase chain reaction
PFGE   pulsed-field gel electrophoresis
RAPD   randomly amplified polymorphic DNA analysis
SNARE   soluble N-ethylmaleimide sensitive factor attachment protein receptor
SVA   Swedish National Veterinary Institute
TeTx   *Clostridium tetani* toxin
VAMP   vesicle-associated membrane protein

1 Foreword

"Fågeldöden kan sprida livsfarlig sjukdom"
[*Transl.* Bird deaths can spread deadly disease]

Those words were the heading of an article hanging on the wall in the laboratory where I started studying botulism. It was in the midst of the tumult caused by the increasing number of avian botulism outbreaks in Sweden. Along with the outbreaks there was a flow of samples coming into the laboratory for analysis. All day long I cut pieces of chicken caeca, cultivated them and analysed them for the presence of *Clostridium botulinum*. It was a tedious job not to mention the very badly smelling working material! However, I did not know it at the time, but these foul smelling chicken samples later became the foundation of this thesis. The number of reported animal botulism outbreaks has increased during the last decade, not only in Sweden, but in several European countries (Skarin *et al.*, 2013). As *C. botulinum* produces an extremely toxic neurotoxin, botulism research became a part of an EU project, AniBioThreat, which aimed to improve the EU's capacity to counter biological animal bio-threats in terms of awareness, prevention and contingency (AniBioThreat, 2010). This opened up the possibility for a PhD project about botulism. Now after four years of research, this thesis represents the content of that PhD project.

2   Background

2.1   *Clostridium botulinum* and botulism

2.1.1 A historical glance

The first mentioning of botulinum poisoning was in Germany in the 17th century, where patients who had eaten blood sausages suffered from paralytic symptoms. The word botulism comes from the Latin word for sausage, botulus. However, the bacterium itself was not mentioned until 1897 in Belgium, where several people became ill and died from eating incompletely salted ham at a wake. A professor in Ghent named Èmile van Ermengem examined the ham in a microscope and discovered rod-shaped bacteria. He called them *Bacillus botulinus*, having in mind the previous reports from Germany where similar symptoms had been described. Van Ermengem established that the illness of botulism resulted from an extremely potent toxin that was produced by this bacterium (van Ermengem, 1979). In 1917 the genus *Bacillus* was divided into *Bacillus* for aerobic species and *Clostridium* for anaerobic species, thus *Bacillus botulinus* was changed into *Clostridium botulinum* (Winslow *et al.*, 1917).

2.1.2 Etiology

Clostridium botulinum is a Gram-positive, anaerobic, rod-shaped bacterium, which occurs as single rods or in short chains (Figure 1). The rods measure 0.3-1.9 um in width and 1.6-9.4 um in length (Smith and Hobbs, 1974). They have flagella and form oval endospores with a varying resistance to heat (from very high to low), depending on the strain (Table 2) (Ito *et al.*, 1967, Segner and Schmidt, 1971, Peck *et al.*, 2011). These spores are distributed in soil and aquatic sediments around the world (Hauschild, 1989, Hatheway, 1990). *C. botulinum* produces extremely potent botulinum neurotoxins (BoNTs), probably with the purpose to kill the host by intoxication. It can propagate

13

from nutrients made available by saprophytic organisms and produce spores and toxin in the anaerobic environment provided by the dead host, thereby ensuring its own spread through larvae or scavenging animals. BoNTs cause the disease botulism affecting both humans and animals, although it is more common in animals (Critchley, 1991). The mortality is high if treatment with antitoxins and respiratory support is not given (WHO, 2013). Animals, which are seldom treated, are occasionally victims to large botulism outbreaks.
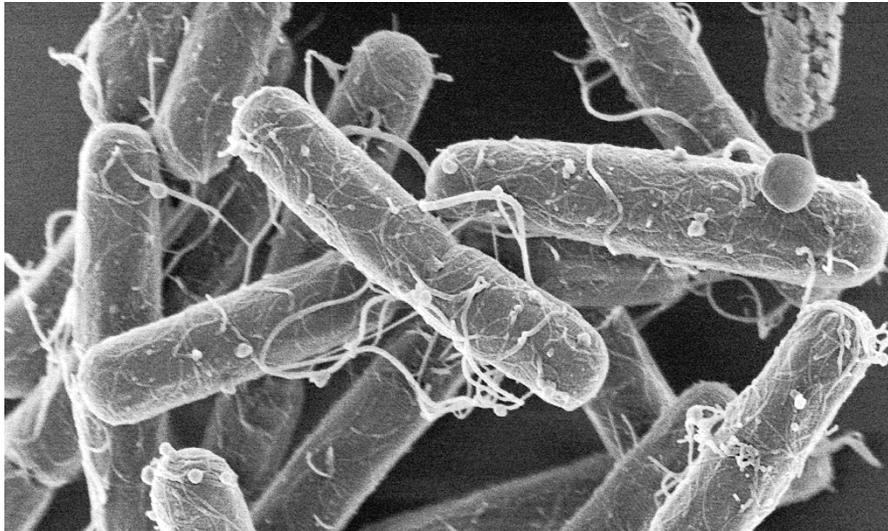
Figure 1. Scanning electron microscopy image of *C. botulinum* vegetative cells (Photo: H. Skarin, SVA and Leif Ljung, Uppsala University)

2.1.3 Taxonomy and classification

C. botulinum is a taxonomic designation for bacteria that produce one or more BoNTs (Smith and Hobbs, 1974). In the beginning of the 20th century it was noted that different strains of *C. botulinum* produced toxins that had different antigenicities, since antitoxins raised towards one toxin did not necessarily neutralize toxins produced by other strains (Leuchs, 1910). These antigenetically distinguishable BoNTs were given alphabetical designations upon discovery; type A and B (Burke, 1919), C (Bengtson, 1922), D (Meyer and Gunnison, 1928), E (Gunnison *et al.*, 1936), F (Möller and Scheibel, 1960), G (Gimenez and Ciccarelli, 1970) and recently H (Barash and Arnon, 2014), although the latter is being questioned (Maslanka *et al.*, 2015). The BoNTs have a high degree of primary sequence conservation but pair-wise comparisons between the seven neurotoxin amino acid sequences (A-G) reveal identities of 30.4-62.8% (Hill and Smith, 2013). *C. botulinum* strains are defined upon which BoNT they produce, thus a *C. botulinum* type A strain

produces BoNT/A *etc*. BoNT/A and B are responsible for the majority of human botulism cases whereas BoNT/C and D are associated with disease in animals. BoNT/E and F are mainly connected to disease in humans, although type E also cause botulism in fish and fish-eating birds (Yule *et al.*, 2006). BoNT/G has so far only once been associated with human botulism (Sonnabend *et al.*, 1981). The most commonly affected species for each BoNT type are presented in Table 1.

Table 1. *Species most commonly affected by botulism and responsible botulinum neurotoxin (BoNT) type*

BoNT type	Most affected species
A	Humans
B	Humans, cattle, horses
C	Birds, cattle, sheep, horses, mink
D	Cattle, sheep
E	Humans, fish, birds
F	Humans
G	Humans (rare)

There are also strains capable of producing two toxins (bivalent strains): type Ab, Ba, Af, Bf, Bh, and AB. When one of the toxins is produced in a smaller amount, it is indicated by a lower case letter. Some strains encode two toxins but only form one active toxin (e.g. A(B) strains). Also, variants (subtypes) within BoNT/A, B, E, and F have been identified, and designated with the type followed by an Arabic number. Recently a strain with three different neurotoxin genes was discovered, one type A and two type F subtypes, and it produces all three neurotoxins (Kalb *et al.*, 2014). The subtypes were originally delineated on the basis of immunological or functional differences, but nowadays more on genetic differences (Smith *et al.*, 2005, Hill *et al.*, 2007, Carter *et al.*, 2009, Raphael *et al.*, 2010). Two mosaic variants of type C and D toxin (type C/D and D/C) have been described, but they are not regarded as additional subtypes (Moriishi *et al.*, 1996a, Moriishi *et al.*, 1996b).

2.1.4 The botulinum neurotoxin

BoNTs are on one hand the most poisonous toxins known and the causative agent of botulism and on the other hand a commonly used drug applied for neurological movement disorders and in the cosmetic industry as an anti-wrinkle remedy (Gill, 1982, Kaji, 2011). BoNT/A became the first biological toxin to receive approval for the treatment of human disease (Johnson, 1999).

BoNT is a large heat-sensitive protein of approximately 150 kDa, consisting of a 50 kDa light chain (LC) and a 100 kDa heavy chain (HC) covalently linked by a disulfide bridge (Schiavo *et al.*, 2000). The HC is in its turn composed of two domains: the N-terminal domain (HC_N) and the C-terminal domain (HC_C). The BoNT-encoding gene (*bont*) is part of a gene cluster, which includes several accessory genes encoding non-toxic neurotoxin-associated proteins, NAPs. Two main *bont* cluster organizations are known, the hemagglutinin (*ha*) cluster and the *orfX* cluster. Both include a gene coding for a non-toxic non-hemagglutinin protein (NTNH), which is more conserved than the *bont* gene. NTNH has a structure highly similar to that of BoNT, and may have evolved from a gene duplication of a common ancestor (Inui *et al.*, 2012). The NTNH directly interacts with BoNT, thereby shielding sensitive parts of BoNT to protect it against the low pH and proteases present in the digestive tract (Niwa *et al.*, 2007). Besides the *ntnh* and *bont* genes, the *ha* cluster typically consists of three hemagglutinin genes encoding Hemagglutinin (HA) and a *botR* gene, whereas the *orfX* cluster contains genes encoding proteins of unknown function and sometimes *botR*. The *botR* encodes a sigma factor 70, which has been suggested to regulate expression of genes in the neurotoxin cluster (Sebaihia *et al.*, 2007).

2.1.5 The disease mechanism

BoNT causes a state of paralysis, and when this happens involuntarily it is called botulism. The disease mechanism is illustrated in Figure 2. BoNT does not penetrate intact skin but may be absorbed via mucosal surfaces or skin lesions (inhalation botulism and wound botulism). In some cases, BoNT is produced upon germination in the body (toxicoinfection). This can happen to infants (infant botulism) and more rarely to immunocompromised patients, but is more common in animal botulism and will be further described later in this chapter. However, the primary route for disease is caused by ingestion of pre-formed BoNT through food or feed. This is sometimes referred to as classic botulism, or food-borne botulism. Depending on disease route, BoNT uses different entries into the blood stream in order to use the circulation to reach nerve endings. In classic botulism, the HA component facilitates adhesion to the intestinal epithelial cells and a small amount of BoNT, alone or in complex with NAPs, is transported across the epithelium. Next, the HA component disrupts tight or adherent junction proteins of the basolateral surface or exerts cytotoxic effects on the epithelial cells, which allows for passage of a larger amount of BoNT or BoNT complexes (Fujinaga *et al.*, 2009, Jin *et al.*, 2009).

BoNT dissociates from NTNH upon entry to the bloodstream where the pH value is above 7 (Gu *et al.*, 2012). It is then transported through the blood to

neurons where it binds to gangliosides, which are enriched at nerve terminals, ensuring a high toxin-cell encounter rate (Rummel *et al.*, 2004, Tsukamoto *et al.*, 2005). Depolarization in the nerve cell triggered by the ganglioside binding initiates vesicle fusion to the plasma membrane and exposes synaptic vesicle proteins, which function as specific co-receptors for BoNTs (specific for HC_C) (Dong *et al.*, 2003, Dong *et al.*, 2006, Dong *et al.*, 2008, Fu *et al.*, 2009, Rummel *et al.*, 2009, Strotmeier *et al.*, 2010, Willjes *et al.*, 2013). Upon binding, the BoNT-receptor complex is transferred into the lumen of the vesicle, and subsequently into the cytosol (Lalli *et al.*, 1999, Koriazova and Montal, 2002). Once inside the cytosol, the light chain is released from the heavy chain. The light chain is a Zn^{2+}-dependent endopeptidase that cleaves specific peptide bonds within neuoronal SNAREs (soluble N-ethylmaleimide sensitive factor attachment protein receptors). BoNT/A, C and E cleave synaptosomal-associated protein (SNAP-25) at different sites and BoNT/C also cleaves syntaxin (Blasi *et al.*, 1993, Binz *et al.*, 1994, Foran *et al.*, 1996). BoNT/B, D, F and G cleave the vesicle-associated membrane protein (VAMP), each at different sites (Schiavo *et al.*, 1992, Schiavo *et al.*, 1993a, Schiavo *et al.*, 1993b, Schiavo *et al.*, 1994). The disruption of SNAREs inhibits the fusion of acetylcholine vesicles with the neuronal membrane, thereby preventing acetylcholine release. This results in paralysis of the affected muscles.
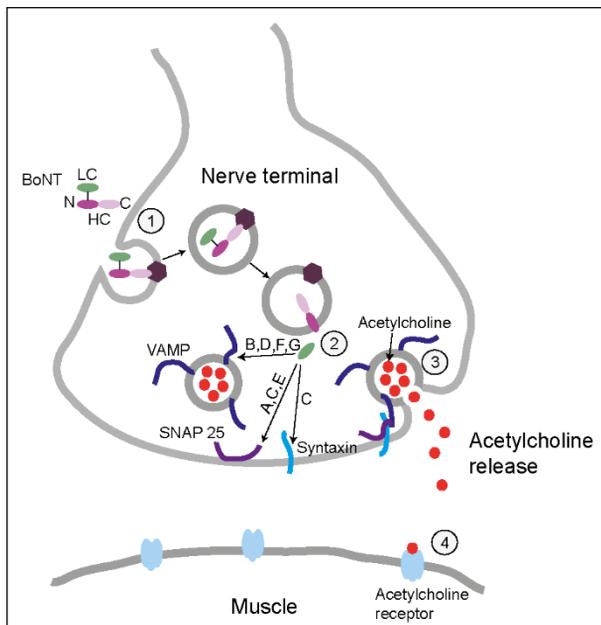
Figure 2. The mechanism of action of botulinum neurotoxins (BoNTs). 1) BoNT binds to a specific receptor and is transferred to the inside of the neuronal cell by endocytosis. 2) The BoNT light chain (LC) dissociates from the heavy chain (HC) and cleaves proteins of the SNARE complex. 3) Without BoNT, acetylcholine is released upon vesicle fusion through SNARE binding with the neuronal membrane. 4) Acetylcholine binding to the acetylcholine receptor on the skeletal muscle fibre is a signal for muscle contraction.

2.1.6 *C. botulinum* and BoNT as a biothreat

Due to the extreme potency, lethality and ease of production, BoNT has been considered a potential bioweapon for decades, either to be used in an aerosol form or by contaminating food or feed (Gill, 1982). There are multiple examples where BoNT has been produced for biological warfare or terrorist attacks, although so far unused or used with unsuccessful result as far as we know (Arnon *et al.*, 2001). Most of these examples rely on aerosolized BoNT, where exposure to the toxin would cause absorption through the lungs. This has been shown experimentally in primates (Franz *et al.*, 1993), but also occurred accidentally to veterinary personnel in a laboratory in Germany (Holzer, 1962). Deliberate contamination of water supplies is generally considered as a minor threat since BoNT is rapidly inactivated by standard potable water treatments (Siegel, 1993). However, the toxin may be stable for several days in untreated water (Siegel, 1993). If food or feed were used as a vehicle for BoNT or *C. botulinum* spores, the outbreak would need to be distinguished from naturally occurring botulism outbreaks in order for actions to be taken against the crime. Both in the US and within the EU there are surveillance systems for human botulism based on clinician reporting, which can aid detecting abnormalities or connections between episodic outbreaks (CDC, 2012, ECDC, 2014). However, animal botulism is not listed as a notifiable disease everywhere and therefore no similar system has been developed, thus making animal husbandry more vulnerable (Skarin *et al.*, 2013).

Botulism research is also an issue of Dual Use Research of Concern (DURC), which is defined by the World Health Organization as research that is intended for benefit, but which might easily be misapplied to do harm. Upon the discovery and suggestion of a new BoNT type (BoNT/H), the sequence was withholded from publication since no antitoxins had yet been developed to counteract the novel toxin (Hooper and Hirsch, 2014). However, it can be claimed that the hold on information is a greater threat than the sequence itself.

2.2 Genomics

The new sequencing platforms have provided an opportunity to gain a deeper understanding of the phylogeny within *C. botulinum*, and to give us new insights regarding the evolution and distribution of virulence factors.

2.2.1 Bacterial genome sequencing

For decades, Sanger sequencing was the golden standard sequencing technology, and it was used to complete the first bacterial genome,

Haemophilus influenza in 1995 (Fleischmann *et al.*, 1995). In later years, there has been a rapid development of next generation sequencing technologies (NGS), which relies on the opportunity to simultaneously read sequences. This results in a considerably higher throughput and deeper sequence coverage. The lower cost per sequenced base in combination with a faster turnaround-time are the main reasons why massive parallel sequencing has caused such a revolution in genome sequencing. Three technologies have been dominating for bacterial genome sequencing: pyrosequencing (454) (Margulies *et al.*, 2005), reversible-end sequencing (Illumina) (Bentley *et al.*, 2008) and semiconductor sequencing (Ion Torrent) (Rothberg *et al.*, 2011). Although based on different chemistries, all require the setup of a DNA library, which is formed by ligating platform-specific synthetic DNA sequences (adapters) onto the ends of the DNA fragments to be sequenced. Next, the library fragments (the template) are amplified by a polymerase-mediated reaction and the template is sequenced in parallel reactions. In the 454 and Ion Torrent instruments, the amplification of template is performed by emulsion PCR onto beads. The beads are loaded onto wells that can fit a single DNA-coated bead each. In 454 pyrosequencing, nucleotides are added one at the time and the incorporation is detected by the light obtained from the release of phosphate molecules. In semiconductor sequencing, an ion sensor detects hydrogen atoms as they are released when a nucleotide binds in to the complementary strand. The light (in 454) or electronic signal (in Ion Torrent) is proportional to the incorporated number of nucleotides. In the Illumina technology, the template is amplified by bridge-amplification directly in the flow cell. All four fluorescently labeled nucleotides are added in each cycle, but they are chemically blocked, which ensures the incorporation of only one nucleotide per fragment cluster in each reaction cycle. This enables the identification of single bases as they are introduced into DNA strands. As the 454 technology is on its way out of the market, a new long-read technology from Pacific Biosciences (PacBio) is uprising. PacBio is a single molecule real time (SMRT) sequencing technique, which by providing reads longer than the longest repeat in a microbial genome can result in an assembly of a single contig (Koren *et al.*, 2013).

The fragmenting of DNA to be sequenced by NGS results in random sequencing start points, which has the advantage that the sequencing output consists of overlapping sequences. These can be combined into a longer sequence, i.e. a contiguous sequence (contig). There are areas in the genome which will result in sequences of poor quality. Further, areas with repetitive sequences and genes existing in multiple copies in the genome, such as ribosomal RNA and transposons, cause additional problems as they result in sequences that can be combined in several ways. Therefore, in order to cover

and combine sequences of a genome into a complete genome (i.e. *de novo* sequencing), additional bioinformatics processing and Sanger sequencing is usually required. A genome that is not fully covered or still consists of multiple contigs is defined as a draft genome.

2.2.2 *C. botulinum* genomics

Already biochemical analysis and later 16S rRNA sequencing showed that *C. botulinum* encompasses genetically diverse strains (Figure 3) (Lee and Riemann, 1970, Cato *et al.*, 1982, Nakamura *et al.*, 1983, Hill *et al.*, 2007). Some strains are capable of digesting complex peptides and are identified as proteolytic, whereas others are non-proteolytic. The growth temperature requirements also differ between strains and so does the spore sensitivity to heat and chemicals. With the increased amount of genetic information about *C. botulinum* strains, it was eventually proposed to divide *C. botulinum* into four different physiological and genotypic groups, I-IV (Table 2).

Table 2. *Characteristics of C. botulinum groups I-IV (Smith and Hobbs, 1974, Miyazaki and Sakaguchi, 1978, Lindstrom and Korkeala, 2006, Carter and Peck, 2015)*

	Group I	Group II	Group III	Group IV
BoNT subtype	A1-A5, B1-B3, F1-F5	B4, E1-E3, E6-E9, F6	C, D, (C/D, D/C)	G
Proteolytic capability	Yes	No	Mixed	Weak
Optimum growth temperature (°C)	35-37	26-30	30-42	30-37
Spore heat resistance	High	Low	Intermediate	Intermediate
Location of *bont* gene	Chromosome/ plasmid	Chromosome/ plasmid	Bacteriophage	Plasmid

In recent years, strains of additional *Clostridium* species have been found capable of producing botulinum neurotoxins, although they have kept their original species designations: *Clostridium butyricum* and *Clostridium baratii* (Hall *et al.*, 1985, Aureli *et al.*, 1986). These strains, which can produce BoNT/E (subtype E4 and E5) or BoNT/F (subtype F7), respectively, have been suggested to be classified as *C. botulinum* group V and VI. As the amount of genetic information from clostridia species has increased, the different groups of *C. botulinum* have been discovered to be more closely related to other clostridia species than to remaining *C. botulinum* groups (Figure 3). For example, *C. botulinum* group I is closely related to *C. sporogenes* and *C.*

botulinum group III to *C. novyi* and *C. haemolyticum* (Sasaki *et al.*, 2001, Hill *et al.*, 2007).
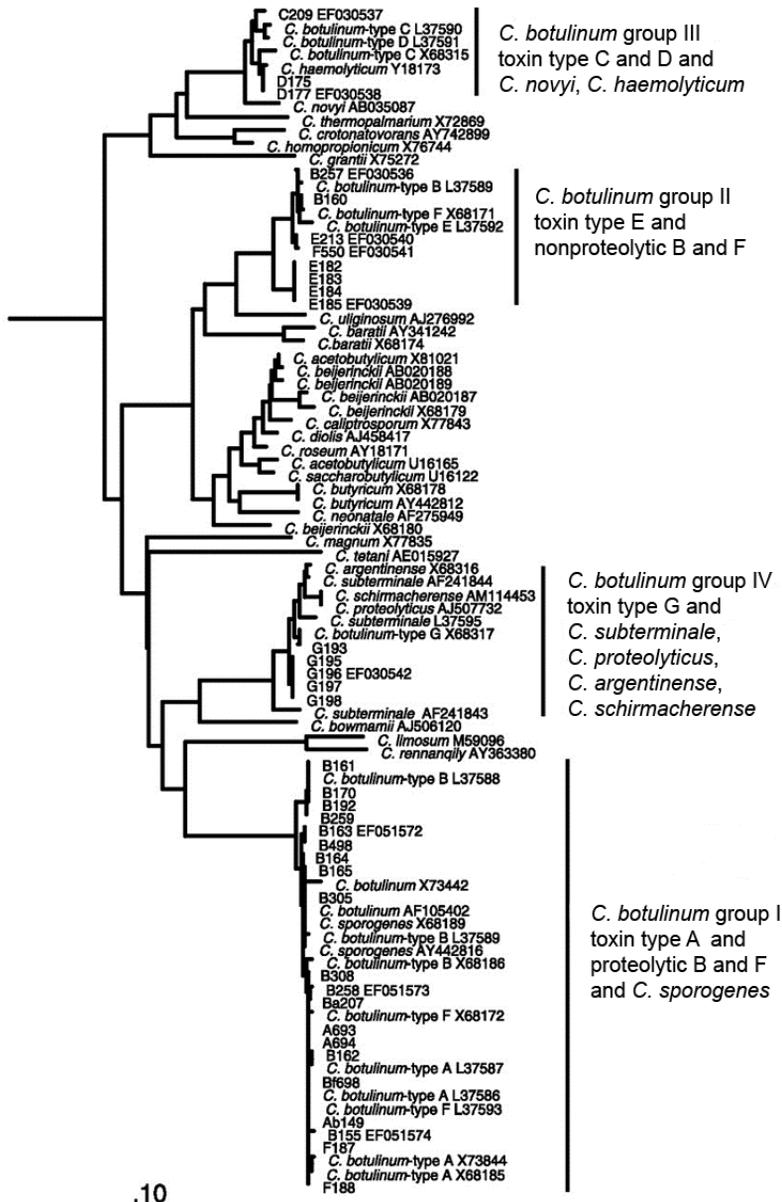

Figure 3. Phylogenetic dendrogram of *Clostridium* species illustrating the genetic diversity of *C. botulinum*. The neighbour-joining tree was constructed using an alignment of 1,329 bp of 16S rRNA gene sequences. Modified from Hill *et al*. (Hill *et al.*, 2007) with permission.

The first completed whole genome sequence from *C. botulinum* was published in 2007 and was from a Sanger sequenced type A strain, ATCC 3502 (Sebaihia *et al.*, 2007). Additional sequencing projects for strains representing other types of *C. botulinum* were soon initiated and to date there are 81 *C. botulinum* genomes available in GenBank, out of which 19 are completed genomes (GenBank, 2015). The genomes so far sequenced of group I strains demonstrate a distinct genetic lineage when compared to genomes of the other *C. botulinum* groups and display only minor insertions/deletions when compared to each other (Carter and Peck, 2015). There are fewer completed group II genomes available for analysis, but the diversity between group II genomes is larger than what is displayed in group I (Carter and Peck, 2015). In group II, genomes of type B and F strains are highly related, whereas type E strains form a separate genetic linage upon whole genome analysis (Stringer *et al.*, 2013). Prior to the projects described in this thesis, there were no complete genomes of *C. botulinum* group III or IV available for analysis.

2.2.3 Mobile genetic elements and horizontal gene transfer

Transfer of DNA between cells is called horizontal gene transfer, and the mobilization is usually mediated by mobile genetic elements such as plasmids, bacteriophages, and transposable elements. This genetic transfer can have a big impact on bacterial genomes as it may affect fitness and pathogenicity by adding or removing accessory genes. Clostridia pathogens are well known for their ability to produce powerful toxins, e.g. the *Clostridium tetani* toxin (TeTx), the *Clostridium perfringens* epsilon toxin, the *C. novyi* alpha-toxin and the *C. botulinum* BoNT. Thanks to whole genome sequencing of clostridia genomes, the locations of these toxin genes have been identified. Several are located on plasmids (*tetx* and the *C. perfringens* epsilon toxin gene) or bacteriophages (*C. novyi* alpha-toxin). The *bont* cluster is organized in different positions on the chromosome or on different plasmids in *C. botulinum* group I and II strains, on large plasmid-like prophages in group III strains, and on plasmids in group IV strains (Table 2). The difference in location of *bont*, in combination with the high overall genetic diversity of the physiological groups, has led to the suggestion that *bont* has been horizontally transferred between different species of bacteria. On the basis of sequence and structural similarities, Collins and East suggested that there was a single ancestral form to both BoNT and TeTx (Collins and East, 1998). Assuming a similar rate of nucleotide change, they further proposed that the ancestral toxin ended up in three or four clostridia subpopulations, which later became *C. tetani*, *C. botulinum* group III, and *C. botulinum* group I and II (or group I and II independently). The clustering of group I strains is less correlated with

neurotoxin type than the clustering of group II strains, suggesting a more frequent horizontal transfer of the *bont* cluster between strains of group I than of group II (Carter and Peck, 2015).

There are several mechanisms that most likely have played parts in the neurotoxin gene transfer and altering events, including conjugation, transduction, transposition and recombination. Conjugation has been demonstrated by a BoNT-encoding plasmid (pCLL) from a *C. botulinum* group II strain to a nontoxigenic *C. botulinum* group I strain, thus proposing a bridge between these two genetically diverse groups (Marshall *et al.*, 2010). Further, the BoNT-encoding phages of *C. botulinum* group III have been demonstrated capable of infecting non-toxigenic strains of either type C or D strains (Eklund and Poysky, 1974).

Different complete or truncated transposons have been discovered around or within the *bont* cluster, suggesting a role in the event of transfer (Hill *et al.*, 2009). Transposable elements use different mechanisms for translocation, for example an RNA intermediate (retrotransposons), a phage y-type mechanism (conjugative transposons) or a "cut and paste" mechanism (insertion sequence elements). Insertion sequence (IS) elements are short DNA sequences (<2,5kb) mainly encoding genes (recombinases/transposases) needed for their own excision and insertion (Mahillon and Chandler, 1998). Among the clostridia species, the *C. difficile* genome contains a large proportion (11% of its genome) of transposable elements, which mainly constitutes of conjugative transposons (Sebaihia *et al.*, 2006), and IS elements are abundant in some *C. perfringens* genomes (Myers *et al.*, 2006). To the contrary, transposable elements are scarce and mostly non-functional in genomes of *C. acetobutylicum* (Nolling *et al.*, 2001), *C. tetani* (Bruggemann *et al.*, 2003) and *C. botulinum* group I (Sebaihia *et al.*, 2007).

Mosaic *bont* genes appear to be products of recombination events, for example the chimeric type C/D and D/C *bont* genes in *C. botulinum* group III strains (Moriishi *et al.*, 1996b). Further, BoNT/A2, BoNT/E6 and BoNT/F5 also appear to be hybrids, which include toxin sequences of other *bont* subtypes and/or unique parts (Chen *et al.*, 2007, Raphael *et al.*, 2010, Hill and Smith, 2013). Last, the *ntnh* gene that precedes *bont* has also been identified as a region of recombination between sequences from different *bont* clusters, for example the hybrid *ntnh* consisting of one part from *ntnh of* type C and one part from type A (Kubota *et al.*, 1996).

2.3 Animal Botulism

The diagnosis of animal botulism is often based on clinical signs and seldom laboratory confirmed, which leads to uncertainty of the prevalence of the disease and of responsible BoNT type. However, among reported cases of animal botulism, BoNT/C, D and the mosaic types C/D and D/C, are dominating. Other BoNT types can also cause botulism in animals, for example type A and B cause botulism in cattle and horses, and type E causes botulism in fish and fish-eating birds (Table 1) (Yule *et al.*, 2006, Johnson *et al.*, 2010, Lindstrom *et al.*, 2010, Johnson *et al.*, 2012). The variation of BoNT type causing animal botulism depends to some extent on the distribution of the spores. For example, *C. botulinum* type E spores are mainly found in aquatic environments, whereas type B spores are more commonly found in soil (Huss, 1980). The geographical presence of the different types in soil also vary (Ortiz and Smith, 1994). There are also significant differences in susceptibility towards the different BoNT types. For example, avian species are much more sensitive to type C/D than to type C or D (Miyazaki and Sakaguchi, 1978, Takeda *et al.*, 2005). Further, bovine neuronal cells seem to be more sensitive to type D/C than to type C or D (Nakamura *et al.*, 2012). The difference in susceptibility can to some extent be explained by the varying abilities of the different BoNTs to pass through the digestive tract in different species. It has for example been shown that the HA component binding to adhesion proteins (E-cadherins) is species specific. The HA of BoNT/C does not interact with human E-cadherin and the HA of BoNT/B does not interact with chicken E-cadherin, both consistent with the epidemiology of botulism (Sugawara *et al.*, 2010).

2.3.1 Disease routes

As in humans, many cases of animal botulism are caused by ingestion of pre-formed toxin, i.e. the disease is due to an intoxication. Animals may directly ingest toxin-contaminated material, such as water or feed, or they may ingest toxins through the consumption of zooplankton or invertebrates that have consumed toxic material. Bovine and equine botulism are often related to the presence of carcasses or litter of poultry or wild birds, in combination with poorly preserved silage (Lindstrom *et al.*, 2010).

In toxicoinfection, there is a ceacal colonization of *C. botulinum* resulting in an *in situ* production of BoNT. In avian species, botulism is often caused by the subsequent digestion of caecal droppings containing high concentrations of BoNT (Miyazaki and Sakaguchi, 1978, Hyun and Sakaguchi, 1989). Toxicoinfection can also cause disease in young animals, such as foals (shaker foal syndrome) (Swerczek, 1980). It has further been proposed to be

responsible for paralytic symptoms called "visceral botulism" or "equine grass sickness" in cattle and horses, respectively, but this has not been conclusively confirmed (Hunter *et al.*, 1999, Bohnel *et al.*, 2001, Seyboldt *et al.*, 2015). Spores are not uncommonly found in fecal samples of healthy animals, indicating that there are certain criteria necessary for establishing a toxicoinfection (Notermans *et al.*, 1981, Dahlenborg *et al.*, 2001, Blomqvist *et al.*, 2009). These animals, and especially birds, may spread *C. botulinum* spores from epizootic areas to previously uncontaminated areas. Remains of poultry or poultry manure have been suspected sources of botulinum spores or BoNT in many animal botulism outbreaks (Hogg *et al.*, 1990, Lindstrom *et al.*, 2004, Livesey *et al.*, 2004).

2.3.2 Avian botulism

The first described cases of avian botulism were referred to as "Limberneck disease" or "Western Duck Sickness", as affected birds get a hanging neck and early cases were discovered in the Western part of the United States (Windingstad and Laitman, 1988). In 1930 it was concluded to be a form of botulism (Giltner and Couch, 1930), and since then avian botulism has been diagnosed in wild birds on all the continents, except the Antarctic (Jensen and Price, 1987). Especially big and recurring outbreaks in wild birds have been reported from the Great Lakes region in the United States (Lan Chun *et al.*, 2015), from Incheon in South Korea (Shin *et al.*, 2010), from wetlands in central Spain (Figure 4A) (Vidal *et al.*, 2013), and from the southeast coast of Sweden (Neimanis *et al.*, 2007). Although the outbreaks around the Great Lakes tend to be caused by BoNT/E, avian botulism is most commonly caused by type C or C/D. Avian botulism also occur in domestic birds and there are numerous reports of botulism outbreaks in poultry from around the world. The first confirmed botulism outbreak in broiler chickens in Sweden was in 2003, and since then there have been several outbreaks every year. The highest number of outbreaks in Sweden occurred in 2008 when 13 different broiler farms were affected (Blomqvist *et al.*, 2009). A botulism outbreak was also recently described in commercial laying hens in Sweden (Figure 4B) (Skarin *et al.*, 2015).
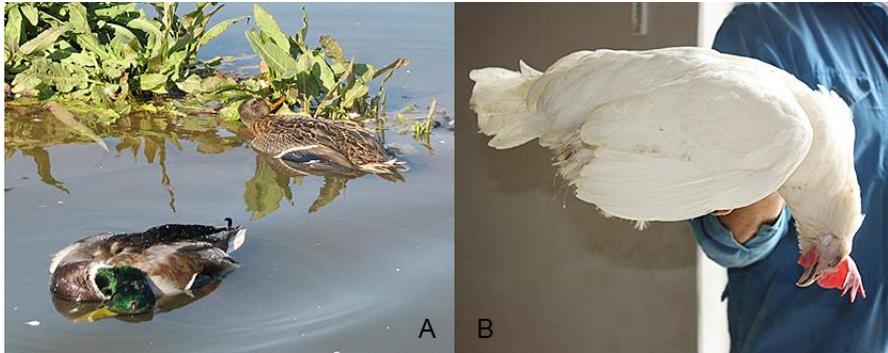
Figure 4. Birds paralysed with botulism. A) Mallards, Spain (Photo: Rafael Mateo, IREC) B) Laying hen, Sweden (Photo: Désirée Jansson, SVA).

2.3.3 *C. botulinum* group III virulence factors

The bont-phage

That BoNT/C toxin production was governed by bacteriophages was suggested already in 1970 (Inoue and Iida, 1970). Soon after, the first phages were isolated from *C. botulinum* type C strains. It was demonstrated that nontoxigenic strains can be easily derived by treating cultures with UV irradiation (Eklund *et al.*, 1971). Hariharan and Mitchell examined supernatants of *C. botulinum* type C strains by electron microscopy and found three different types of tailed phages: a larger, a smaller and an intermediate type with a head of the smaller size but with a tail of the larger size (Hariharan and Mitchell, 1976). Mainly the large, but also the intermediate phage, have been connected to BoNT production. The toxin is produced during the lysogenic state and the balance between lytic and lysogenic state seems to be the main reason to the varying degree of toxicity between strains (Hariharan and Mitchell, 1976).   What affects this balance is not known, but environmental factors are probably involved. For example, the *bont*-phage is more prone to start a lysogenic relationship with the bacterium at 30°C than at 37°C. Further, less dense cultures are infected at a higher rate than older cultures (Eklund *et al.*, 1987).

Phages from different *C. botulinum* group III strains, capable of producing either type C or D toxin, have been demonstrated to be able to infect nontoxigenic strains, even if the strain originally was able to produce the other toxin type. However, some strains are insensitive to phages from another strain, or become only weakly toxic (Eklund and Poysky, 1974). Also, alpha-toxin encoding phages isolated from *C. novyi* type A supernatants are able to infect and produce toxin in both toxigenic and nontoxigenic *C. botulinum* type C and D strains (Eklund *et al.*, 1974).

The *bont*-phage is not integrated into the chromosome but resides in the cytoplasm as a large circular prophage during the lysogenic state (Sakaguchi *et al.*, 2005). The discovery of such phage-like plasmids are emerging with the number of genomes being analysed, and they are described for example also in *Streptomyces* (Zhong *et al.*, 2010) and *Salmonella* (Octavia *et al.*, 2015).

The C2 and C3 toxins

In addition to BoNT, two ribosylating exoenzymes, the C2 and C3 toxin, have been described in *C. botulinum*, both being produced solely by *C. botulinum* group III strains. Experiments in animals have shown that they exhibit enterotoxic properties, but their role in animal botulism is still not clear (Ohishi, I., 1983). The C2 toxin is composed of two polypeptide chains: one is the binding component and the other is the enzymatic component, which catalyzes the ADP-ribosylation of nonmuscular G actin (Aktories *et al.*, 1986). The C3 toxin acts on the Rho protein (Aktories *et al.*, 1989). The genes encoding these two toxins are carried on mobile elements. The binary C2 toxin genes are plasmid-borne, and the gene encoding the C3 toxin is located on the *bont*-phage (Popoff *et al.*, 1990, Sakaguchi *et al.*, 2009). Further, the C3 toxin gene has been proposed to be located on a transposon of 21.5kb, as its surrounding sequence bares similarities with site-specific transposons and since this 21.5 kb sequence is more conserved than the rest of the phage (Hauser *et al.*, 1993). Also, this 21.5kb fragment was missing in a phage mutant obtained by nitrosoguanidine treatment (Oguma *et al.*, 1975, Hauser *et al.*, 1993).

2.3.4 Diagnostics

The diagnosis of animal botulism is often made by the veterinarian on the basis of clinical signs and the absence of diagnostic postmortem findings. For laboratory confirmation, it is considered essential to detect active botulinum neurotoxin, and the golden standard method is still the mouse bioassay (Tevell Aberg *et al.*, 2013). In this method, samples to be tested are injected into mice, which are observed for symptoms of botulism. The toxin type can also be determined by injecting the sample along with an antitoxin and observe the mice for symptoms of botulism. The advantage by this method, which has made replacement difficult, is the high sensitivity (1 mouse lethal dose is approximately 10 pg) (Grate *et al.*, 2010). However, in the past decade, several new methods aiming to replace the mouse bioassay have been developed. Some of those are based on the protease activity of BoNTs, for example mass spectrometric detection of the BoNT cleavage products (Hedeland *et al.*, 2011, Kalb *et al.*, 2015).

Detection of bont type C, D or mosaic C/D and D/C genes

Besides detecting the toxin, there are also different Polymerase Chain Reaction (PCR) methods developed for detection of *bont*. PCR is a method used for amplification of a sequence, which allows for DNA detection by staining (conventional PCR) or emitted fluorescence as the DNA is being amplified (real-time PCR). A cultivation step to enrich for anaerobic spore-formers is usually required prior to *bont* detection by PCR (Lindstrom and Korkeala, 2006). The first PCR method developed for a *C. botulinum* group III strain was developed in 1992 for the *bont*/D gene (Sunagawa *et al.*, 1992), followed by *bont*/C (Szabo *et al.*, 1993). Now there are many conventional and real-time PCR methods developed for detection of *bont* genes in *C. botulinum* group III strains. The first PCR method that was able to differentiate between type C, D and the mosaic types C/D and D/C genes was developed by Takeda *et al.* and was a conventional PCR (Takeda *et al.*, 2005). Several multiplex real-time PCR methods have been developed within the AniBioThreat project to fasten and simplify the diagnostics of animal botulism by being able to differentiate between type C and D, or type C/D and D/C (Woudstra *et al.*, 2012) or between all four *C. botulinum* group III *bont* types (Anniballi *et al.*, 2013).

Isolation and genotyping methods

The isolation of *C. botulinum* group III is generally known to be difficult, partly due to the strict anaerobic metabolism and lack of known selective traits, but also because of the instability of the *bont*-phage, since *bont* detection often is used for species confirmation. The epidemiology of strains causing animal botulism is to a large extent unknown, mainly because of the difficulty to isolate pure cultures of *C. botulinum* group III.

Today different targeted PCR-based approaches for genotyping of bacteria are common, e.g. Multilocus sequence typing (MLST) (Perez-Losada *et al.*, 2013) and Multiple-Locus Variable number tandem repeat Analysis (MLVA) (Lindstedt, 2005). However, the epidemiological studies performed so far concerning animal botulism outbreaks have applied more traditional genotyping techniques: pulsed-field gel electrophoresis (PFGE), amplified fragment length polymorphism (AFLP), and randomly amplified polymorphic DNA analysis (RAPD) (Grenda *et al.*, 2014). The main advantage with those techniques is that they can detect genetic variations in the whole genome without any prior knowledge of the genetics of the organism. In PFGE, genomic DNA is fragmented by a restriction enzyme and size distributed by applying an electric field that constantly changes direction, which allows for separation of very large fragments (Figure 5). AFLP is a PCR-based method, in which restriction fragments of genomic DNA are amplified prior to separation

and comparison, usually by automated capillary sequencing instruments. RAPD is another PCR-based method, where a low annealing temperature and only one non-specific short primer (usually 10 bp) is applied to achieve random amplification.
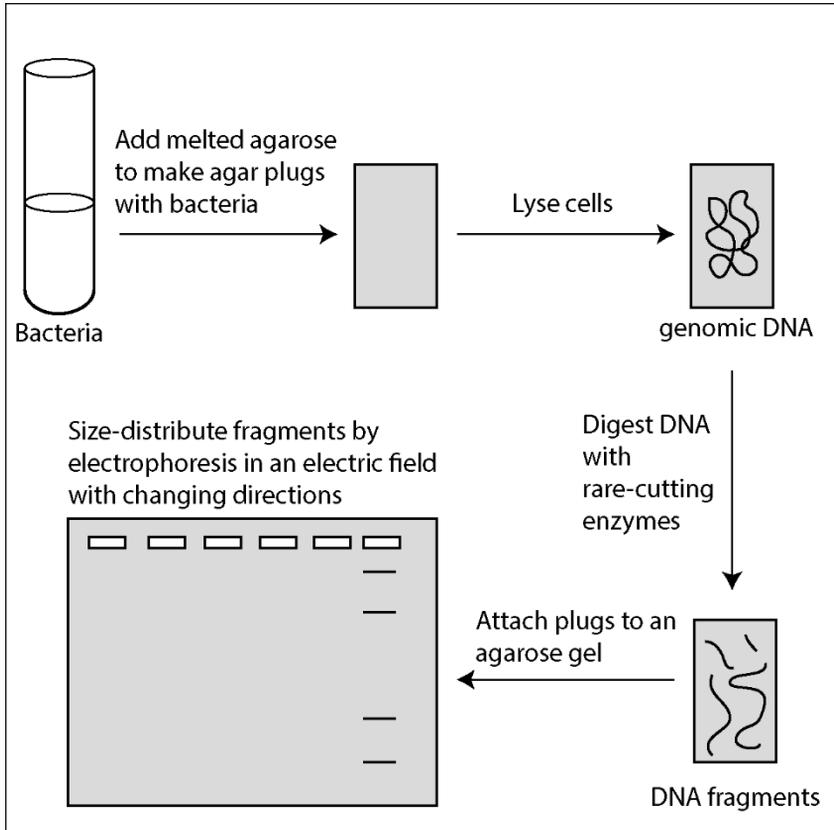
Figure 5. Schematic description of pulsed-field gel electrophoresis (PFGE).

3   Aims of the thesis

This thesis project was part of the European collaboration project AniBioThreat. The main objective of the thesis was to increase the knowledge about *C. botulinum* group III strains causing animal botulism, and especially avian botulism, with a special focus on European strains. More specifically, the aims were:

- ➢ To develop methods for isolation and molecular characterization of *C. botulinum* group III strains
- ➢ To isolate and characterize *C. botulinum* group III strains from animal samples
- ➢ To investigate the genomic diversity of *C. botulinum* group III strains
- ➢ To study the genomic organization of *C. botulinum* group III strains
- ➢ To explore the genomic relationship between *C. botulinum* group III, *C. novyi* and *C. haemolyticum*

The first aims to develop methods for isolation and characterization of *C. botulinum* group III strains and the application of these methods are mainly covered by paper I and II. However, the strains collected during this thesis project are also the basis for paper III and IV. To investigate the genomic diversity of *C. botulinum* group III strains causing animal botulism has in one way or the other been the aim in all papers included in this thesis. The *C. botulinum* group III genomic organization has been investigated in detail in paper III and IV. The genomic (and more specifically plasmidomic) relationship between *C. botulinum* group III, *C. novyi* and *C. haemolyticum* is brought up already in paper I and III, but is thoroughly investigated in paper IV.

4 Comments on Materials and Methods

4.1 Isolates

4.1.1 Sampling materials

Many of the *C. botulinum* strains isolated in this thesis project are isolated from avian samples from the vast number of avian botulism outbreaks in Sweden (both in wild birds and poultry) since 2003. Some isolation material from animal botulism outbreaks was received from other partners of the AniBioThreat consortium. Both the Norwegian and the Danish strains were isolated from samples sent to The National Veterinary Institute (SVA) for *bont* detection analysis. The South Korean strain was isolated from a pig liver sample by the author during a visit at the Animal and Plant Quarantine Agency in South Korea. The Spanish strains, obtained from avian samples collected at three different wetlands in central Spain, were isolated at SVA in collaboration with the Spanish PhD student Ibone Anza Gomez. Most of the *C. novyi* and *C. haemolyticum* strains used in this study, were either obtained from culture collections or received from other researchers. One *C. novyi* strain was isolated from a sample collected at a botulism outbreak in broiler chicken in Sweden.

4.1.2 Cultivation

Broth cultivation procedures were adopted from previous studies at SVA, but the composition of the tryptose-peptone-glucose-yeast extract (TPGY) broth had to be adjusted. Sodium thioglycolate was originally used as oxygen reducing agent, but as the *C. botulinum* strains suddenly stopped growing in broth and this compound is known to sometimes inhibit growth, the reducing compound was changed into a mixture of cysteine hydrochloride (cysteine-HCl) and sodium bicarbonate ($NaHCO_3$) (Segner *et al.*, 1971). Similarly, the McClung Toabe agar (McClung and Toabe, 1947) used for cultivation on agar plates was also supplemented with cysteine-HCl and $NaHCO_3$. Normally 1 g of

sample material was inoculated into 9 ml of pre-reduced broth and this was put in a water-bath at 70°C for 15 min to kill vegetative cells. The sample was then incubated for 2 days at 37°C. *C. botulinum* group III strains can grow in a broad temperature range, from 10 to 47°C (Mitchell and Rosendal, 1987). A lower incubation temperature (30°C) has been used for marine strains (Segner *et al.*, 1971) whereas a higher temperature (42°C) may be relevant for avian strains reflecting the body temperature of birds. The temperature of 37°C was chosen to provide for growth of all kinds of strains connected to animal botulism. Incubation was either done inside an anaerobic cabinet or in anaerobic jars.

4.1.3 Isolation

The largest challenge during the isolation process was the competing flora and it was therefore found necessary to pre-heat the samples in order to select for spore-formers. Cultures were prepared as in 4.1.2 and processed after a minimum of 24 hours to ensure the presence of spores. Initially isolates were obtained by many re-streaks to obtain single and pure colonies. This was often unsuccessful. Instead, antibodies were produced (by Innovagen) towards spores of six different *C. botulinum* type C and C/D strains. A portion of spore-containing material was mixed with antibodies labeled with biotin. The spore-antibody complex was then coupled to streptavidin-coated magnetic beads and extracted from the sample using magnetic force. The procedure is further described in paper II. As antibodies were produced against complete spores, they are likely not specific for *C. botulinum* spores, but rather bind to spore coat proteins also in spores from other bacteria. However, this immunomagnetic separation (IMS) technique was found valuable to decrease competing flora and provided for isolation of the majority of the isolates obtained during this thesis project.

4.1.4 Identification of strains

After learning how to recognize them, *C. botulinum* group III colonies were usually easily detected on the McClung agar containing egg yolk, on the basis of colony morphology. Especially due to the iridescent sheen due to lipase activity and an opaque precipitate surrounding the colony due to lecithinase activity. However, the colonies from a pure culture sometimes had a mixed morphology: they could be smooth or have a sprawling appearance with or without the surrounding precipitation (Figure 6A). *C. botulinum* group III strains have mixed proteolytic capabilities and this may cause a difference in colony morphology (Smith and Hobbs, 1974). This was not confirmed as these strains were not analysed for proteolytic capability, however it seems unlikely

that different colonies from the same strain would have different proteolytic capability. It was established that the difference in morphology did not correlate with loss of the *bont*-phage. It has been claimed that by using the McClung egg yolk agar it should be possible to distinguish between *C. botulinum* group III strains and *C. novyi* type B and *C. haemolyticum* strains (McClung and Toabe, 1947). Considering the mixed morphology sometimes present in a culture, it was found impossible to make such a distinction.

   C. botulinum group III strains growing in TPGY broth sink to the bottom of the tube as they reach stationary growth phase and they do not leave a bad smell (Figure 6B). Turbid growth and a foul smell were always signs of contamination in an overnight culture.
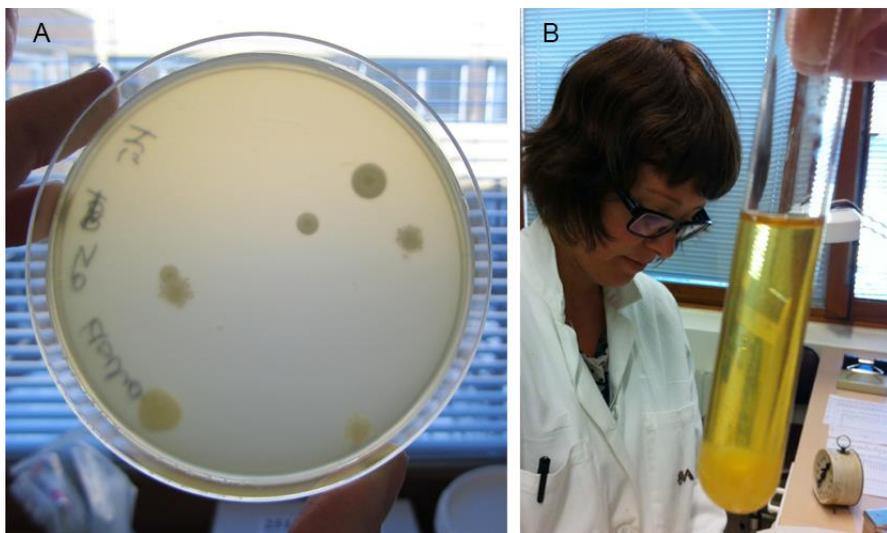
Figure 6. Cultures of *C. botulinum* group III. A) Colonies of mixed morphology on McClung Toabe agar. B) Culture in tryptose-peptone-glucose-yeast extract broth.

C. botulinum and *C. novyi* strains were primarily identified upon the presence of *bont* and the *C. novyi* alpha-toxin gene, respectively. The PCR protocol mainly used for *bont* detection in this thesis detects *bontC* and *bontC/D* (Lindberg *et al.*, 2010). The two *C. haemolyticum* strains used in the study were type strains. Since it was sometimes difficult to identify strains on the basis of toxin genes due to phage loss during the isolation process, a chromosomal PCR was designed to be used in addition to toxin gene profiling. This PCR (described in paper II) specifically detects the gene encoding the 50S ribosomal protein L10 of *C. botulinum* group III, *C. novyi* and *C. haemolyticum*.

4.2   DNA extraction

DNA of some clostridia species has been noted to be easily degraded by endogenous DNases (Kristjansson *et al.*, 1994). *C. botulinum* is clearly one of those species, as purified genomic DNA run on an agarose gel typically results in a smear of degraded DNA. This is usually not a big problem for DNA to be used in a PCR analysis as it does not require long pieces of DNA. For this purpose, DNA was extracted either by boiling, by the DNeasy blood-and-tissue kit (Qiagen, Germany), in the BioRobot EZ1 workstation (Qiagen) or in the case of a large number of samples by using Magattract DNA mini M48 kit (Qiagen) in the Magnatrix automated nucleic acid extraction machine (NorDiag). However, whole genome sequencing required a higher concentration of high quality DNA. The DNA extraction kit that generated the best result for this purpose was the MasterPure Gram Positive DNA Purification Kit (Epicentre Biotechnologies, USA).

4.3   Genotyping

PFGE is considered the primary technique for genotyping *C. botulinum* strains as it is applied in various laboratories throughout the world, and standardization (i.e. through PulseNet) has enabled comparisons between different laboratories. PFGE was chosen as the main genotyping technique in this thesis project for several reasons: required equipment was available, it is functional for inter-laboratory comparability of generated results, and it does not require any previous sequence information. However, PulseNet protocols recommended for *C. botulinum* group I or II strains (CDC, 2013) could not be applied for group III strains. The cell lysis buffer and ES buffer were basically the same, except mutanolysin, which had no effect on group III strains, was removed. The main adjustments to the protocol were the fixation of cells in formaldehyde, which decreased endogenous DNase activity, and the use of HEPES buffer instead of a Tris-containing buffer for electrophoresis. Without those two changes the bands were barely visible or only a smear. This is shown in Figure 1 in paper I. Still not all strains were typeable by PFGE. AFLP, which requires much smaller DNA fragments than PFGE, might have been a safer choice for genotyping. However, banding patterns generated by AFLP are also affected by poor DNA quality (Vuylsteke *et al.*, 2007).

Two restriction enzymes, *Sma*I and *Sal*I, were chosen for PFGE restriction analysis on the basis of the number of bands produced and their discriminatory ability. Since *Sma*I is more commonly used for other PFGE analysis, it was continuously used for genotyping of *C. botulinum* group III strains. Between 7

and 15 bands (25-407 kbp) were produced by PFGE of the analysed *C. botulinum* group III strains.

A protocol for RAPD was also optimized for *C. botulinum* group III strains in order to have a faster genotyping method than PFGE available (one day compared to five days). The RAPD method was as discriminatory and reproducible as the PFGE analysis when only including strong bands. However, there were many weaker bands generated, which made the results more difficult to interpret.

4.4   Sequencing and assembly

4.4.1 Sequencing technology

The main advantage with 454 sequences to be used for *de novo* assembly is the read length (<1kb for GS FLX Titanium) (Roche, 2015), which overruns the ones generated by the Illumina (250 bp for HiSeq and 250 or 300 bp for MiSeq) (Illumina, 2015) and Ion Torrent instruments (400 bp) (LifeTechnologies, 2015). The PacBio RS II platform generate the longest sequences, with an average read length of 10-15 kb (PacBio, 2015). However, the latter technology could not be used in this project due to the requirements of unfragmented DNA. Therefore the first genomes that were sequenced in these studies were sequenced with 454 sequencing at the Science for Life Laboratory (Stockholm, Sweden). Remaining genomes were sequenced on Illumina MiSeq (2x250 bp) because of the lower price and easier access (an instrument was installed at SVA). Besides, the MiSeq sequences provide paired-end information, which is useful for the assembly program to accurately assemble sequences into contigs. The high sequence throughput generated by Illumina HiSeq was used to correct errors in completed plasmids.

4.4.2 Genome assembly

All genomes were *de novo* assembled as the abundance of IS sequences and variations in the genomes made reference-guided assembly unsuitable. However, the sequence of the chromosome from *C. novyi*-NT (GenBank accession CP00382) was used for guidance when ordering contigs in the *C. botulinum* BKT015925 genome. The generated 454 sequences were assembled using GS assembler (Newbler, Roche applied science). Sequences generated by MiSeq were assembled using two different open source software packages: MIRA (Chevreux *et al.*, 2004) and SPAdes (Bankevich *et al.*, 2012). SPAdes was found to be the best choice of assembly program for MiSeq sequences from *C. botulinum* group III strains, resulting in the lowest number of contigs.

Contigs were further processed and assembled using Consed (Gordon *et al.*, 1998).

4.4.3 Plasmid assembly

First, uncleaved genomic DNA was separated on an agarose gel by PFGE in order to get an estimate of the number and size of plasmids in each strain. Originally S1 nuclease was used to linearize the plasmids but this step was removed as the DNA was sufficiently nicked by endogenous DNAses to be able to migrate in the gel. Different approaches were used in order to identify sequences belonging to plasmids. First, plasmid DNA was separated from genomic DNA by using plasmid purification kits: plasmid mini kit (Qiagen) and NucleoBond PC20 (Macherey-Nagel, Germany). However, this resulted in a DNA concentration too low for sequencing. Instead, strains were cured from plasmids by a combination of heat and novobiocin treatment (McHugh and Swartz, 1977). Sequences from the cured strains were compared to sequences from corresponding wild type strains and differing sequences were identified as plasmid DNA.

4.4.4 Annotation

Open reading frames (ORFs) in the completed genome and plasmid sequences were identified with Glimmer (Delcher *et al.*, 1999). Well-conserved genes were automatically annotated from other clostridia genomes and previously annotated *C. botulinum* group III plasmids. Automatic annotations were provided on the basis of both identity (>60%) and coverage (>90%) of match to annotated reference sequences. Less certain annotations were manually assigned by protein alignments using the basic local alignment sequence tool (BLAST) against the NCBI non-redundant protein sequences database. Protein sequences were further analysed for conserved domains by SwissProt.

4.5 Genome comparative analysis

The Gegenees software (Agren *et al.*, 2012), which uses a fragmented approach to make all-against-all BLASTN whole genome comparisons, was used to perform the phylogenomic analysis of genomes and plasmids. For whole genome comparative analysis, the average normalized BLAST scores from comparisons of all 200 bp fragments could be used as a measurement of overall genomic similarity. Upon plasmid comparative analysis, a threshold of 40% was used to define the relative amount of conserved sequences (as fragments falling under that threshold are not included in the average similarity

value). The phylogenomic overviews were illustrated by heat-plots (such as Figure 7B).

MUMmer (Kurtz *et al.*, 2004) was used for genome and plasmid comparisons and the Mauve Genome Alignment software (Darling *et al.*, 2010) and the Artemis Comparison Tool (Carver *et al.*, 2005) were used for plasmid alignments.

4.6 Plasmid analysis

A PCR system detecting the different plasmids characterized in paper III was used to screen for plasmids in closely related unsequenced strains. Three or four primer pairs were chosen for each plasmid: two targeting plasmid replication/partitioning genes, and one or two targeting putative virulence genes. Detection of genes was performed using PerfeCTa SYBR® Green SuperMix Low ROX (Quanta BioSciences) and real-time PCR as described in paper II. Conjugation was investigated using the ClosTron mutagenesis system (Heap *et al.*, 2007). The p3BKT015925 plasmid was tagged with a group II intron, ensuing erythromycin resistance, and mixed with a *C. novyi* type A strain (07-BKT029909). Conjugants should be able to form colonies in the presence of erythromycin (2.5µg/ml) and nalidixic acid (8µg/ml).

5 Results and Discussion

5.1 Collection and identification of strains (paper I, II and IV)

To be able to study the genomic organization and diversity of *C. botulinum* group III it was of importance to have a representative strain collection. Before the work of this thesis project begun, there was only one *C. botulinum* group III strain in the SVA culture collection of bacterial strains, the C-Stockholm strain, which was isolated from mink in 1949 when SVA was located in Stockholm (Dinter and Kull, 1951). This strain is now used as a reference strain by many laboratories around the world. However, as *C. botulinum* group III strains are known to be difficult to isolate, this was the first challenge.

Most of the isolates obtained in this work originate from avian samples, partly because the majority of the samples were from avian botulism outbreaks, but also because birds diseased from botulism more often than other animals have had an *in situ* proliferation of *C. botulinum*, and therefore are more likely to contain the bacteria. From the avian samples, it soon became evident that it was much easier to isolate pure cultures from liver samples than from caecum samples due to the low amount of contaminating flora present in the liver. However, not all the birds diseased from botulism and with *C. botulinum* present in the caecum had spores in the liver. The presence of competing flora was the largest problem when isolating pure cultures since *C. botulinum* was easily overgrown by other better growing species. It was therefore a big advantage to work with samples that had been collected shortly after death and preserved cold. The scarcity of such well-kept animal samples from cases with a diagnosis of botulism contributed to the challenge of isolation. Pre-heating and eventually also IMS targeting *C. botulinum* spores were found to be beneficial to obtain samples with as little competing microbes as possible. Another problem that arose during the isolation of strains was the inability to identify strains when they had lost the phage carrying the *bont* gene. The more

difficult samples, which required several re-streaks, had an increased risk of losing the phage. It was noticed that some strains lost the phage easily, whereas others, such as the C-Stockholm strain, have kept the phage for many generations. We know that the lytic cycle of the phage to some extent is triggered by environmental factors, but it is not known how the lysogenic relationship with the host differs between strains.

In total, 42 *C. botulinum* group III strains were obtained during the work of this thesis, the majority being of Swedish avian origin. But the collection also comprised avian strains from Norway, Denmark and Spain, bovine strains from Italy and The Netherlands, and one porcine strain from South Korea. All the avian strains isolated were of type C/D and two of the three bovine strains were of type D/C. In a related study performed during the work of this thesis it was discovered that among samples naturally contaminated with *C. botulinum* spores collected in France, Italy and the Netherlands, only chimeric types could be detected (Woudstra *et al.*, 2012). The majority of samples positive for *bontC/D* originated from avian species, and *bontD/C* from bovine species, but there were also some cases of vice versa. These chimeric types were first described in the 1990's (Moriishi *et al.*, 1996a, Moriishi *et al.*, 1996b) and today it seems that they are much more commonly involved in botulism outbreaks than type C or D strains. This is probably to some extent an effect of the improved diagnostics that differ between type C, D and mosaic forms, but may also be because they are more toxic to their host than either type C or D (Takeda *et al.*, 2005, Nakamura *et al.*, 2012).

5.2   Genomic diversity in *C. botulinum* group III (paper I, II, III and IV)

Genomic diversity plays an important role in the adaptability and survival of a species. However, *C. botulinum* dormant spores have little need for adaptation, and strains of this species are generally associated with a living host only during short periods, which may explain the relatively stable genomes of *C. botulinum* group I and II strains. At the onset of the work of this thesis, very little was known about the genomic diversity in *C. botulinum* group III.

This was studied by genotyping the strain collection with PFGE. Thirty-seven strains were successfully genotyped and resulted in the pulsotypes demonstrated in Figure 7A. A subset of pulsotypes, which accounted for 27 of the strains, clustered closely together and differed by only one or two bands. Remaining pulsotypes were represented by a single or a few strains. Twelve *C. botulinum* group III genomes chosen to represent a diversity of both pulsotype and origin were sequenced (Figure 7). Pair-wise comparisons of the average

sequence similarity of these genomes and two older *C. botulinum* group III draft genomes retrieved from GenBank formed four lineages. Lineage I mainly consisted of strains with chimeric *bont* genes (type C/D and D/C) of both avian and bovine origin. Lineage I was designated two subgroups: Ia and Ib, on the basis of the very high sequence similarity between strains within these subgroups. Subgroup Ia corresponds to the largest cluster of related pulsotypes and Ib to the second largest pulsotype cluster (Figure 7). Lineage II comprised *C. botulinum* type C or D strains isolated from mammals. Lineage III and IV were diverged from lineage I and II and consisted of strains of avian or bovine origin encoding chimeric BoNT types. Care should be taken not to interpret two strains with very different PFGE banding patterns as genetically distantly related. For example, comparison of the PFGE banding patterns from strains of the two subgroups in lineage I revealed a very low similarity. Thus, genotyping of *C. botulinum* group III by PFGE identifies closely related strains but tells less about the overall genomic similarity between strains.
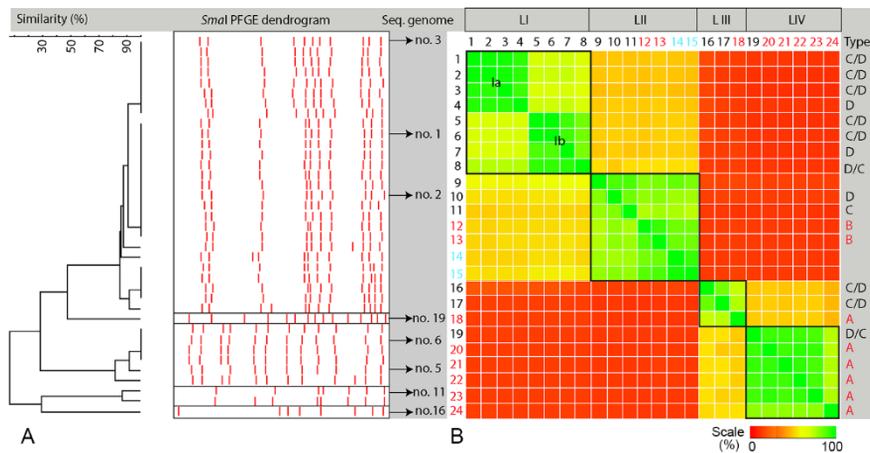
Figure 7. An overview of genotyped and whole genome sequenced strains. A) Pulsed-field gel electrophoresis *Sma*I dendrogram of genotyped *C. botulinum* group III strains. Pulsotype clusters are framed B) Heat-plot illustrating pair-wise comparisons of the average sequence similarity of *C. botulinum* group III genomes (black), *C. novyi* genomes (red), and *C. haemolyticum* genomes (blue). Genome nr 1 is the completed genome described in paper III (BKT015925). Lineages (LI-LIV) are framed.

After the publication of paper IV, there was another genotyping study published using a flagellin detection assay targeting four different flagellin C gene types (*fliC*-I to *fliC*-IV) identified in strains of *C. botulinum* group III (Woudstra *et al.*, 2015). This assay was applied on 204 samples naturally contaminated with *C. botulinum* group III strains originating from five different European countries. All *bontC/D*-positive samples contained the *fliC*-

I type, which by far was the most commonly detected type (n=172). The reference sequence used for the *fliC*-I type was from BKT150925, which corresponds to one of the strains in the largest pulsotype cluster in Figure 7A and belongs to genomic lineage I (subgroup Ia). There is also an AFLP-analysis previously published, where genotyping of 25 *C. botulinum* group III strains resulted in only four clusters (Hill *et al.*, 2007). The C-Stockholm strain was used, which made it possible to identify one of the clusters as lineage II recognized in this thesis project. That cluster also contained strains of non-chimeric types (C and D) and were mainly from mammals. Interestingly, another cluster (with only two strains) clustered closest with *C. botulinum* type G strains.

Although more strains need to be analysed, this data suggests that *C. botulinum* group III can be divided up in only a few genomic groups containing strains of widely diverse origins. The overall pair-wise similarity of the most distantly related *C. botulinum* group III strains is lower than that seen in *C. botulinum* group I or II (data not shown), which thus can be considered to be less diverged than group III. The reason for this may be that the *bont*-phage present in *C. botulinum* group III has a broader host-spectrum than the *bont*-carrying plasmids present in *C. botulinum* group I and II. It can also be speculated that the longer periods of *C. botulinum* group III strains propagating inside a host, which is the result of toxicoinfection, contribute to a higher rate of genome alterations than that seen in *C. botulinum* group I and II.

5.3   The *C. botulinum* group III genome (paper III)

One of the type C/D strains (08-BKT015925) from the largest pulsotype cluster (Figure 7) was chosen for whole genome sequencing and assembled into completion. The strain was isolated from a sample collected from a botulism outbreak in broiler chickens in Sweden 2008. Its genome consists of a circular chromosome and five circular plasmids, all with a low GC content of 26-28%. A schematic image of the BKT015925 genome is presented in figure 1 in paper III. Putative functions could be predicted for 66% of the chromosomal coding sequences (CDS) and for 38% of the total CDS in the five plasmids. The BKT015925 genome is smaller than the group I and II genomes, 3.2 Mbp compared to 3.6-4.4 Mbp. Furthermore, the number of plasmids in BKT015925 was found to be much higher than in the other two *C. botulinum* groups, where no strain has been characterised with more than two plasmids. The five plasmids in BKT015925 were named with designated prefixes of p1-p5 ordered in decreased size. Only p2BKT015925 and p3BKT015925 were thought to be regular plasmids and the remaining three were considered to be

phages due to the presence of genes predicted to be associated with phage functions. However, these phage sequences are circular and also contain genes involved in plasmid replication and/or partitioning, thus they seem be prophages propagating like plasmids. The presence of several smaller phages within one strain, besides the *bont*-phage (p1BKT015925), also correlates with early electron microscopy examinations of growth culture supernatants from *C. botulinum* group III strains (Hariharan and Mitchell, 1976). As these plasmid-like phages are difficult to categorise, they will from here on be included in the general term plasmids.

Another discovery in the *C. botulinum* group III genome was the abundance of insertion sequence (IS) elements. Five different IS elements (ISCbo1-5) were found in a high copy-number and in different positions in closely related strains, indicating that they are actively re-locating. All together, these mobile elements account for 3.4 % of the BKT015925 chromosome, and they are also located on the three largest plasmids. Eight additional re-locating IS elements, IsCbo7-13 (IS200/IS605 elements), were identified on the two largest plasmids, but they appear to be less active. A higher density of IS elements was found on the two regular plasmids (p2BKT015925 and p3BKT015925) than on the ones considered to be phages. Generally, transposons are considered to be a disadvantage to phage functionality. The IS elements were found to be inserted in non-coding regions on both replichores, except ISCbo2, which had a tendency to be incorporated into another mobile element, a tandem reverse transcriptase (ISCbo3). Some of these elements are related to previously identified IS elements in other bacteria species (Table 3), suggesting that horizontal gene transfer events have taken place between those species. ISCbo1, ISCbo4 and elements related to IS200/IS605 appear also in *C. botulinum* group I or II strains, although mainly in truncated form, in regions flanking the *bont* cluster. This implies that these elements have played some part in transfer of *bont*. However, only one IS element (ISCbo1) was located with close proximity to the *bont* cluster in one of the group III genomes. This was a complete element, as opposed to the same element found in a truncated form next to *botR* in *C. botulinum* group I *OrfX* clusters. The absence of traces of IS elements in the *bont* cluster in *C. botulinum* group III strains in combination with the smaller genome size could indicate a higher pressure to remove non-essential DNA.

Table 3. *Actively re-locating Insertion Sequence (IS) elements identified in the BKT015925 genome and their relation to other IS elements.*

IS element	IS family	Positions in BKT015925	Presence in *C. botulinum* group I or II	Related element
ISCbo1	IS6	28	Yes (truncated)	ISCpe7 (*C. perfringens*)
ISCbo2	IS21	14	Yes (truncated)	IstA and IstB
ISCbo3	-	8		Reverse transcriptase
ISCbo4	IS256	8	Yes (truncated)	ISCpe3 (*C. perfringens*)
ISCbo5	IS1182	18	Yes (truncated)	ISbma2 (*Burkholderia mallei*)
ISCbo6-13	IS200/IS605	9	Yes	OrfA and OrfB

5.4 *C. novyi sensu lato* (paper I, III and IV)

The definitions of the three species *C. botulinum*, *C. novyi* and *C. haemolyticum* are established on the basis of the toxin they produce: the BoNT of *C. botulinum* group III, the alpha-toxin of *C. novyi* type A and B, and the beta toxin of *C. novyi* type B and *C. haemolyticum*. However, they were already before the revolution of sequencing suggested to be closely related on the basis of DNA-DNA hybridization and biological characteristics. As a matter of fact, *C. haemolyticum* has since the late 1950s sometimes been referred to as *C. novyi* type D (Oakley and Warrack, 1959). The discovery that phages of *C. novyi* type A were able to infect *C. botulinum* group III strains, thus phenotypically converting them into *C. novyi*, further contributed to the assumption of a close relationship between these species. This was later established when 16S rRNA sequences were analysed from strains of the different species (Sasaki *et al.*, 2001).

Early in these studies, a *C. botulinum* type C strain (C-Stockholm) was concluded to be more closely related to *C. haemolyticum* and *C. novyi* type B than to the isolated avian type C/D strains, on the basis of 16S sequence comparisons. Later, the completed sequence of the chromosome from *C. novyi* strain NT was a valuable resource for defining the contig order when completing the genome sequence of the *C. botulinum* type C/D BKT015925 strain. There were non-conserved regions scattered throughout the sequences, but all together a high conservation of gene synteny between the genomes. Also *C. novyi* and *C. haemolyticum* genomes were sequenced and compared to *C. botulinum* group III genomes. It turned out that the *C. novyi* and *C. haemolyticum* genomes could be organized into three of the same lineages as

the *C. botulinum* strains (Figure 7). Lineage I was the only one consisting solely of *C. botulinum* strains. The high average sequence similarity seen between genomes in the subgroups in lineage I suggests that they constitute two evolutionary younger branches than remaining lineages. Lineage II comprised strains of all three species (*C. botulinum* type C and D, *C. haemolyticum* and *C. novyi* type B). *C. novyi* type A strains were organized into the relatively diverged branch composed of lineages III and IV.

From these results it is clear that the current nomenclature formed on the basis of the toxin genes is not supported by a phylogenetic concept of a species. A new term, *C. novyi sensu lato* (meaning *C. novyi* in a broader sense), was proposed for a genospecies including *C. botulinum* group III, *C. novyi* and *C. haemolyticum*. Mainly responsible for the phylogenetic inconsistency are the phages harboring the species-defining toxin genes, which are able to move between different strains of different species. This will be discussed in the next two sections.

5.5   The large mobile plasmidome of *C. novyi sensu lato* (paper III and IV)

Compared to the rather conserved chromosomes of *C. novyi* senu lato genomes, the remarkably large plasmidome was found highly variable within the lineages. All together 61 plasmids were identified and 23 of them were assembled into completed sequences. Many of these plasmids were considered to be phage-like plasmids. Apart from randomly distributed related genes, there was no apparent shared synteny between these plasmids and plasmids in other clostridia, including the ones from the other *C. botulinum* groups.

Strains of *C. botulinum* group III, *C. novyi* type B and *C. haemolyticum* contained between three and five plasmids each, whereas strains of *C. novyi* type A only contained one (the *C. novyi* alpha-toxin phage). Genomes in lineage subgroup Ia contained the highest abundance of plasmids (four to five each). A PCR system detecting the different plasmids in lineage Ia was applied on non-sequenced strains belonging to the largest pulsotype cluster. It appeared that they all contained four or five plasmids each, always the regular plasmids corresponding to p2BKT015925 and p3BKT015925 in combination with at least two phage-like plasmids.

The plasmids identified in *C. novyi sensu lato* were organized into 13 plasmid groups (PG) on the basis of the relative amount of shared genetic material, in combination with the conservation of plasmid maintenance genes. An image of the genetic relationship between the plasmids and their organization into plasmid groups is presented in figure 2 in paper IV. Some of

the plasmid groups contained a substantial variety of plasmid origin; from strains belonging to different genomic lineages and from strains of different species. For example, the *bont*-phage (PG1) was found in all lineages (with or without *bont*). The plasmid sequences in most of the plasmid groups mimicked the phylogeny of the whole genome sequences, indicating an infrequent or temporary horizontal movement between strains of these species. The exceptions were the *C. novyi* alpha-toxin phage (PG9 and PG10), which is described further below, and a *bont*-phage (PG1) derived from lineage III but closest resembled *bont*-phages from lineage I. The other *bont*-phages from lineage III and the only one in lineage IV had already been lost in the population upon cultivation. The scarcity of PG1 plasmids in those two lineages suggests that the lysogenic relationship is only temporary in these strains. It is possible that an infection by these phages into a new host does not result in a stable lysogenic relationship, perhaps due to incompatibility with existing plasmids.

The second most dynamic plasmid group was PG10, with plasmids from lineage II, III and IV. The PG10 plasmids from lineage IV contained the gene coding for the *C. novyi* alpha-toxin, whereas plasmids from lineage II and III did not. This plasmid group was genetically related to PG8, PG9 and to a less extent also to PG11, all consisting of phage-like plasmids with a modular structure. Alignment of these plasmids showed a high similarity in some clusters of functionally related genes between plasmids of the different groups, in contrast to the very low similarity surrounding these clusters. This shows a tendency of exchanging modules of functionally related genes with each other, which has resulted in a very mosaic group of phage-like plasmids.

The plasmid groups that consisted of regular plasmids contained no outliers with discordant phylogeny compared to its corresponding genomes, indicating a stable genomic relationship and a low transfer activity. Only two putative conjugation genes were found on these plasmids, but additional genes associated with conjugation can have been missed since many of the plasmid genes lacked predicted functions. A plasmid of a lineage Ia strain (08-BKT015925) was tagged using the ClosTron mutagenesis system (Heap *et al.*, 2007). The modified strain was co-incubated with a *C. novyi* type A strain (07-BKT029909), but no conjugants were detected. It is still possible that conjugation can occur in the environment at certain favorable conditions, but the data collected in this thesis project shows no signs of plasmid transfer between lineages, except when it is governed by phages. In general, the phage-like plasmids found in *C. novyi sensu lato* seem to be more unstable than the regular plasmids as they were easier to remove from a strain by heat or by chemical treatment.

5.6 Pathogenicity in *C. novyi sensu lato* (paper III and IV)

The three species in *C. novyi sensu lato* are known for their ability to cause different diseases: botulism, gas gangrene (*C. novyi*), black disease (*C. novyi*) and bacillary hemoglobinuria (*C. haemolyticum*), respectively. Both black disease and bacillary hemoglobinuria, which affect herbivores, are caused by germination of spores in the liver upon liver damage (usually by liver flukes), which reduces oxygen levels. As already mentioned, these diseases are caused by different toxins. It appeared upon analysis of sequenced *C. novyi sensu lato* genomes that besides the transfer of toxin-encoding phages, toxin genes also show an independent variability of location.

The genes encoding BoNT and the C3 toxin were found on all PG1 plasmids, except the one derived from *C. haemolyticum* (Figure 8). The presence of a PG1 phage without these toxin genes supports the idea that they at some point have been inserted or deleted from PG1. The fact that both toxin genes are missing raises the question if their presence is connected. The mobile element harboring the C3 gene was also missing in a phage with a mutated *bont* gene, further supporting this idea (Hauser *et al.*, 1993). The *bont* cluster sequences were more conserved (all were *ha* clusters) and their locations more consistent in *C. botulinum* group III than in *C. botulinum* group I or II, indicating less mobility of the *bont* cluster or that it has been more recently introduced into *C. botulinum* group III (Figure 8). The only big genetic alterations that seem to have happened to the *bont* cluster in group III strains are the ones responsible for the chimeric sequences *bontC/D* and *bontD/C*. Sequence alignment of chimeric sequences against non-chimeric sequences revealed different recombination sites in the two chimeric sequence types. The estimated recombination sites are illustrated in figure 4 in paper IV.
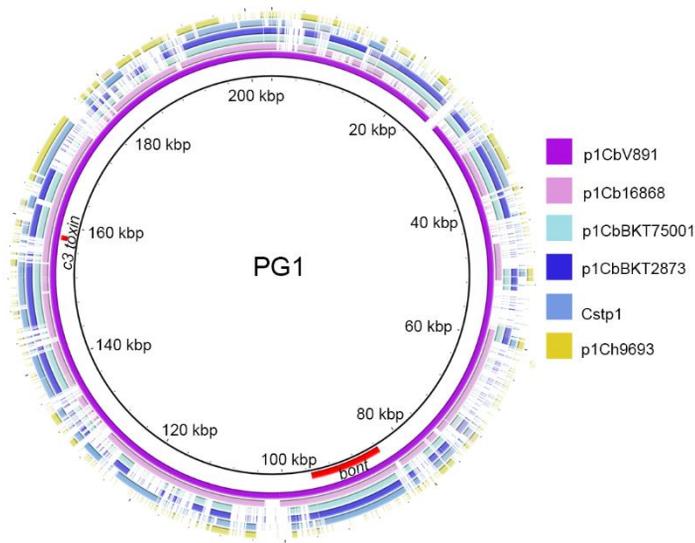
Figure 8. BLAST comparisons of PG1 plasmids from lineage I and II against p1BKT015925. Nucleotide sequence identity > 50% are colored and visualised by BRIG (Alikhan *et al.*, 2011). Note the conserved botulinum neurotoxin gene (*bont)* and the c3 toxin gene, except in p1Ch9693 from *C. haemolyticum*.

The gene encoding the *C. novyi* alpha-toxin was found in all *C. novyi* strains but on different plasmids (PG9 and PG10). The phospholipase C gene, which constitutes the beta toxin in *C. novyi* type B and *C. haemolyticum*, was chromosomally located, except in lineage Ia strains where it was found on a plasmid (PG3). The genes encoding the C2 toxin components were only found in *C. botulinum* group III strains of lineage Ia and II, but were localized on different plasmids (PG3 or PG4).

Besides these more known toxins, other putative toxin candidates are also produced by strains of *C. novyi sensu lato* (Figure 9). Besides the phospholipase C gene, only one other putative toxin gene was found on the chromosome and it was closest related to tetanolysin, which is believed to contribute to the pathogenesis in an infection by *C. tetani*. The high degree of conservation and its existence in all genomes so far analysed of *C. novyi sensu lato,* indicate an important function and also one more evolutionary link to *C. tetani*. The gene coding for an extracellular cysteine endopeptidase of high similarity to clostripain produced by *C. histolyticum* (>75% amino acid identity), was found in all the lineages, but also in *C. botulinum* group I and II strains. This enzyme has been proposed to be the endogenous protease responsible for cleaving the neurotoxin into a light and heavy chain, which is required for its toxicity (Sebaihia *et al.*, 2007). Interestingly, the clostripain gene was located on the chromosome in *C. botulinum* group I and II strains and in *C. novyi sensu lato* lineage III and IV strains, but on PG2 plasmids in lineage I and II strains, adjacent to another putative toxin gene encoding an aerolysin. The aerolysin gene was found in all lineage I and II strains and

shared 53-64% amino acid identities (highest score in lineage I strains) with the alpha-toxin of *C. septicum*. Lastly, two tandem located putative epsilon type B toxin genes were identified on different plasmids of *C. botulinum* strains (PG7 and PG4).

All the identified putative toxin genes, except tetanolysin, had a variable presence or were found on different replichores in different strains (Figure 9).
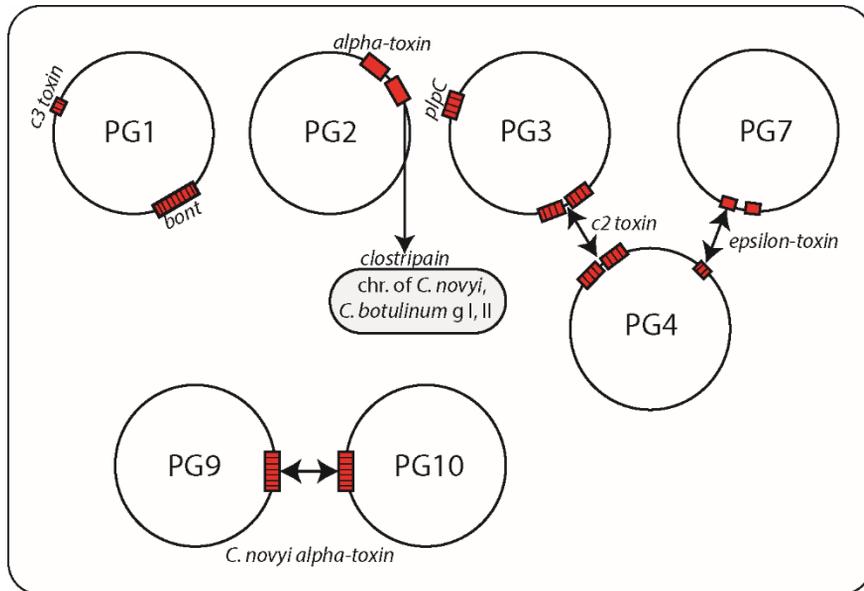
Figure 9. The location of toxin genes in *C. novyi sensu lato*. Toxin genes are illustrated with red boxes. Striped boxes illustrate a varying presence of the toxin gene within the plasmid group. Botulinum neurotoxin gene (*bont*) and Phospholipase C gene (*plpC*).

Transposons were found adjacent to several of the toxin genes, to *bont* (ISCbo1), the *C. novyi* alpha-toxin gene (ISCbo10 and a site specific recombinase), and the epsilon B toxin gene (a site-specific recombinase), suggesting involvement in the transfer. The re-location of toxin genes, by phage movements, by recombination, or by transposons, is likely to change the pathogenicity of the strain and complicates the definition of the species. In this project, one *C. novyi* type A strain was isolated from a chicken caecum sample collected during a botulism outbreak. This shows how entwined these organisms are, and that there are opportunities for the phages to move between strains of different lineages. When a new strain is isolated it will be identified upon the presence of the toxin-encoding phage, even if it was only a temporary phage infection. This is reflected by outliers revealed in phylogeny studies, for example one *C. novyi* strain that clustered with *C. sporogenes* (Stackebrandt *et*

al., 1999), and *C. botulinum* group III strains that clustered with *C. botulinum* type G strains (Hill *et al.*, 2007). All are probably consequences of the dynamic toxin-encoding phages.

5.7   The successful lineage I

The subgroup Ia of lineage I was overrepresented among strains isolated during this thesis project, and among the vast number of European samples analysed by Woudstra *et al*. (Woudstra *et al.*, 2015). Lineage I seems to be the natural host for phages carrying *bontC/D* and *bontD/C*. As previously discussed, the increased toxicity of chimeric *bont* genes may have led to a raised number of botulism outbreaks and thereby the isolation of an increased number of lineage I strains. However, this is probably not the complete answer as there are more strains genotyped corresponding to subgroup Ia than of Ib. Strains in subgroup Ia were the only ones containing all the toxin genes described above, except the *C. novyi* alpha-toxin. Perhaps a combination of the functions of all putative toxin genes contribute to more virulent strains. The highest density of plasmids and IS elements was also found in subgroup Ia genomes, which might have contributed to the evolvement of a better adapted genome, perhaps better adjusted to cause toxicoinfection.

6   Conclusions

When the studies for this thesis started no culturing of clinical samples was applied at SVA. During 2007 and 2008 when the number of avian botulism outbreaks increased by ten-folds, we still did not know much about the causing strains, only that they were *C. botulinum* type C/D. As a result of this thesis project, culturing, molecular identification and genotyping methods for *C. botulinum* group III are now established at SVA. Thanks to these improved and implemented diagnostic methods it was possible to obtain well characterised strains, which were used to study both the organization and diversity of the *C. botulinum* group III genomes.

The *C. botulinum* group III genome is organized into a conserved chromosome with little excessive material, which instead is kept on highly plastic and dynamic plasmids (including phage-like plasmids). This is probably an efficient solution allowing for rapid adaptation when changing habitat, for example from soil to animal intestines. The *C. botulinum* group III strains analysed so far by different genotyping techniques can be organized into only a few genomic lineages, which all contain strains of widely diverse origins. These lineages also comprise *C. novyi* and *C. haemolyticum* genomes, which shows the close genomic relationship between these three species that are collectively described as *C. novyi sensu lato*. The mix of species in the lineages is the result of the dynamic nature of phages carrying the toxin genes upon which these species are identified. Actually, most of the putative toxin genes identified in *C. botulinum* group III are located on plasmids and are also as the many IS elements, actively or sporadically re-locating. Some of the putative toxin genes and IS elements are related to genes found in other bacteria species (mainly clostridia species), which show that *C. botulinum* group III genomes are to some extent influenced by its environment by horizontal gene transfer.

One *C. novyi sensu lato* genomic lineage seems to have developed as *C. botulinum* specific and consists of strains most commonly causing avian botulism, and possibly animal botulism in general, in Europe today. The

majority of the strains in this lineage encode chimeric BoNTs of type C/D or D/C. The spread of these possibly more virulent strains may have contributed to the elevated number of botulism outbreaks being reported during the last decade.

7   Perspectives for the future

Botulism is a devastating and costly disease for the animal husbandry and deserves more attention in the veterinary and scientific field. There are not enough efficient countermeasures or official guidelines developed for handling animal botulism outbreaks. Further, if animal botulism was a notifiable disease in all countries, the number of officially reported cases would increase and give a more accurate picture of the prevalence of the disease. It would also add to an awareness of fluctuations in outbreaks and to epidemiological conclusions.

When writing this thesis some future research needs became clear:

➢ In this thesis project, a number of putative new virulence factors were identified. These need to be further characterised in terms of expression and functionality.

➢ Little is still known of the basic biology of the *bont*-phage. What stabilizes the lysogenic relationship and what triggers the lytic cycle are two questions that are still to be answered. Further, the host-bacteria relationship in toxicoinfection is hardly explored at all. It would be valuable for the animal husbandry to know if there are certain triggers, such as a feed component or the composition of intestinal flora, for toxicoinfection to be established.

➢ What we need now in terms of diagnostics are fingerprinting profiles that can be exchanged by laboratories easier than the first generation genotyping techniques such as PFGE and RAPD. As a result of this thesis there are now a number of different *C. botulinum* group III genome sequences easily accessible to the scientific community. These can be used to define for example MLST loci with a power to discriminate also between closely related strains. Genome sequencing will in the future likely be commonly applied by diagnostic laboratories, but in the meantime an MLST method

can be applied both by the laboratories with access to whole genome sequencing and those using Sanger sequencing or commercial sequencing services.

8   Populärvetenskaplig sammanfattning

Bakgrund

Jordbakterien och sporbildaren *Clostridium botulinum* bildar botulinumtoxin (BoNT), ett av världens farligaste gifter. BoNT är ett nervgift som orsakar förlamningssjukdomen botulism hos både människor och djur. Vanligast är botulism hos djur, som ibland drabbas av stora utbrott med hög dödlighet. I Sverige är det vanligast med botulism hos fjäderfä och framför allt slaktkycklingar har drabbats hårt det senaste decenniet. Det har även rapporterats botulismutbrott bland sjöfåglar i Sverige. Botulism hos fåglar, s.k. aviär botulism, förekommer i hela världen, och stora utbrott har rapporterats de senaste åren från bl.a. Spanien och Sydkorea. Andra djur som drabbas relativt ofta av botulism är nötkreatur, häst, får och mink.

Botulism uppstår oftast av att man får i sig förbildat toxin, vanligtvis genom mat eller foder, s.k. klassisk botulism. För att *C. botulinum* ska kunna växa och bilda gift krävs en totalt syrefri miljö, såsom i konservburkar eller ensilage. I sår kan också en delvis syrefri miljö uppstå där gift kan bildas och orsaka sårbotulism. I mer sällsynta fall kan bakterien växa till sig i tarmen (toxikoinfektion) hos vuxna människor och djur, men denna form av botulism är desto vanligare hos unga djur, spädbarn (spädbarnsbotulism) och fåglar.

C. botulinum är en väldigt divergerad bakterieart som brukar delas in i fyra grupper (grupp I-IV) baserat på genetiskt innehåll och fysiologiska egenskaper så som t.ex. tillväxttemperatur och sporernas resistens mot värme. Gemensamt för dessa är att de alla har förmågan att bilda BoNT. Det har visat sig att stammar av de fyra grupperna egentligen är närmare besläktade med andra bakteriearter än med varandra. T.ex. är *C. botulinum* grupp III nära besläktad med *Clostridium novyi* och *Clostridium haemolyticum*, två arter som orsakar helt andra sjukdomar än botulism.

Hittills har sju olika typer av BoNT konfirmerats (typ A-G) med hjälp av olika antikroppars förmåga att neutralisera toxinets verkan. BoNT typ A, B, E

och F orsakar majoriteten av botulismfallen hos människor. Typ G har än så länge endast vid ett tillfälle påvisats i samband med human botulism. Botulism hos djur orsakas vanligtvis av BoNT typ C och D, även om typ B också är förknippat med botulism hos nötkreatur och häst och typ E med fisk och sjöfåglar. Två hybrider av BoNT typ C och D (typ C/D och D/C) upptäcktes på 90-talet och orsakar idag många av botulismutbrotten hos djur. Det har visats att BoNT typ C/D är mer giftig för fåglar än BoNT typ C eller D, och att BoNT typ D/C troligen är mer giftig för nötkreatur än typ C eller D. BoNT typ C, D, C/D och D/C bildas av bakterier tillhörande *C. botulinum* grupp III. Dessa bakterier bär genen som kodar för toxinet på en bakteriofag (bakterievirus) som ligger vilande i bakterien. Ibland lämnar fagen bakteriecellen och då tappar cellen förmågan att bilda BoNT.

C. botulinum bakterier av grupp III är speciellt svåra att odla fram i renkultur. Dels för att de är mycket syrekänsliga, men framför allt för att de växer sämre än många andra bakterier som också finns i miljöprover och prover från djur. Ett annat problem som ibland uppstår gäller korrekt identifiering av bakterien. Vanligen identifieras den genom att man detekterar BoNT eller genen som kodar för BoNT. Om fagen som bär genen som kodar för BoNT har lämnat bakterierna blir det svårt att identifiera stammen som *C. botulinum*. Dessa problem ligger till grund för att man inte vet så mycket om de stammar som orsakar botulism hos djur.

Studier och Resultat

I studierna som ledde fram till den här avhandlingen utvecklades och förfinades metoder för att isolera stammar av *C. botulinum* grupp III för att kunna studera deras genetiska uppsättning och jämföra dem med varandra. Majoriteten av stammarna som isolerades var från prover som samlats in under svenska utbrott av aviär botulism. Det isolerades också stammar från fågelprover insamlade i Norge, Danmark och Spanien. Några stammar från nötkreatur från Italien och Nederländerna och en stam från gris från Sydkorea ingick också i stamkollektionen.

Det totala DNA-innehållet hos dessa stammar jämfördes med hjälp av pulsfälts-gel elektrofores (PFGE). I PFGE klyvs DNA i olika långa fragment med hjälp av restriktionsenzym som klyver specifikt vid vissa kombinationer av baserna A, T, G och C. De olika långa DNA fragmenten separeras sedan i storleksordning i en gel genom att applicera ett elektriskt fält som konstant byter riktning. På så sätt får man fram ett mönster av band i olika storlekar som kan jämföras med mönster från andra stammar av samma art. Detta visar vilka stammar som är närbesläktade med varandra. Det visade sig att majoriteten av de stammar som orsakat aviär botulism var närbesläktade, oavsett geografisk

härkomst. Dessa stammar var också närbesläktade med en stam isolerad från nötkreatur.

Tolv stammar valdes ut att representera olika pulsfältsmönster, samt härkomst från olika geografiska platser och olika djurslag. Dessa helgenomsekvenserades, dvs. det totala DNA-innehållet i form av korta överlappande DNA-sekvenser lästes av. Denna metod har en högre upplösning än PFGE, eftersom man istället för ett DNA-mönster har tillgång till den totala DNA-sekvensen (genomet). När DNA-innehållet jämfördes kunde stammarna delas upp i fyra olika genetiska undergrupperingar. DNA-sekvenserna i ett av dessa genom, från en stam isolerad från ett svenskt botulismutbrott på slaktkyckling 2008, pusslades ihop till ett komplett genom. Detta studerades närmare och visade sig vara mindre än genom av *C. botulinum* grupp I och II. Det mest intressanta med genomet i övrigt var att det innehöll ovanligt många (fem stycken) mindre självständiga DNA-enheter (plasmider). Tre av dessa plasmider visade sig vara vilande bakterievirus som uppför sig som plasmider (bl.a. den fag som bär genen för BoNT). Vidare innehöll genomet flera potentiella toxingener samt många återkommande (repetetiva) gener och bägge dessa gentyper påträffades på olika ställen vid jämförelse med genom från andra stammar av *C. botulinum* grupp III. Flera av dem liknar motsvarande gener i andra bakteriearter, vilket indikerar att dessa har överförts mellan bakterier. Ibland tar repetetiva gener med sig omkringliggande DNA-sekvenser när de flyttar på sig, t.ex. närliggande toxingener, vilket troligen har hänt vid flera tillfällen i *C. botulinum*. Detta kan påverka bakteriens förmåga att orsaka sjukdom.

Då *C. botulinum* grupp III-genomen jämfördes med genom av *C. novyi* och *C. haemolyticum* visade det sig att de senare var så pass lika några av *C. botulinum*-stammarna att de kunde delas in i samma undergrupperingar. Dessa arter är så närbesläktade att de genetiskt borde anses vara samma art. Fenotypiskt finns ju dock vissa skillnader i och med att de har förmågan att bilda olika toxin och därmed orsaka olika sjukdomar. Den största anledningen till att stammar av dessa arter är så genetiskt ihopblandade i de olika undergrupperingarna är att de fager som bär på toxingener (vilka används för identifikation av både *C. botulinum* grupp III och *C. novyi*) ibland flyttar sig mellan dessa stammar.

Slutsatser

C. botulinum är en divergerad art på grund av att genen som kodar för BoNT har överförts mellan olika mindre närbesläktade stammar. Genen som kodar för BoNT typ C, D, C/D och D/C hos stammar av *C. botulinum* grupp III bärs av en fag som också kan infektera stammar av *C. novyi* och *C. haemolyticum*.

Detta gör att dessa istället identifieras som *C. botulinum*. I och med att dessa tre arter är så närbesläktade vore det mer korrekt att dela in dem i samma art. Av den anledningen har vi föreslagit den kollektiva termen *C. novyi sensu lato*. Dessa studier tyder på att stammar av *C. botulinum* grupp III kan delas in i endast ett fåtal genetiska undergrupper. En av dem består av bara *C. botulinum* stammar och den motsvarar majoriteten av de stammar som isolerats från aviär botulism (oavsett geografiskt usprung), samt majoriteten av de stammar som producerar hybridtyperna av BoNT (C/D och D/C). Detta tyder på att dessa möjligen mer virulenta stammar sprider sig över Europa och är ansvariga för ökningen i antalet fall av botulism hos djur som rapporterats de senaste decennierna.

9   Acknowledgements

Many people have helped and supported me during my PhD studies and I am grateful to you all. Especially I wish to thank:

MY SUPERVISORS

Eva Olsson Engvall, my main supervisor, for being a mentor and protector of my PhD research time. For always being ready to help and for signing a lot of time sheets. But also for excellent reading and travelling advices. I hope you get to do much more of those things soon when I have more time to do some of your work… I look forward to walk in your footsteps!

Bo Segerman, my closest scientific supervisor, I have learned so much from you! Thank you for introducing me to genomics and for all your patience with me and my lack of programming skills. I still owe you a pizza.

Viveca Båverud and **Anna Lindberg**−the original botulinum research group at the department of bacteriology. You are the ones who offered me a job at SVA and taught me about botulism and PCR, and for this I am grateful. Anna, thank you for keeping both feet on the ground and for being such a sensible

person. Viveca, thank you for your encouraging support throughout this project!

Rickard Knutsson, the project coordinator of the AniBioThreat project. Without you this PhD project would not have been possible. It has been three very interesting, inspiring, productive and fun years. When writing this thesis I have tried to remember to keep it CLEAR.

Gunilla Trowald-Wigh, my supervisor at SLU, for administrative support and for being my link and communication channel to BVF.

MY CO-AUTHORS

Anna Aspán, **Gunilla Blomqvist**, **Désirée Jansson**, **Therese Håfström** and **Josefina Westerberg** for valuable scientific input and nice collaborations.

MY COLLEAGUES IN THE ANIBIOTHREAT PROJECT

For the friendly and inspiring atmosphere at our meetings. An extra big thank you to **Camilla Beijer** and **Susanna Westerberg**, for the outstanding project organization and for correcting my time sheets! And to you and **Jeffrey Skiby**, for all the help with the animal botulism workshop. Also, a big thank you to those I have collaborated more closely with in Task 5.2: **Mikael Hedeland**, **Annica Tevell-Åberg**, **Kristian Björnstad**, **Cedric Woudstra**, **Patrick Fach**, **Miriam Koene**, **Luca Bano**, **Fabrizio Anniballi**, **Dario De Medici,** **Charlotta Löfström and Trine Hansen**.

MY COLLEAGUES AT THE DEPARTMENT OF BACTERIOLOGY

For guiding in the labs in "bakt-korridoren", for helping me with SVALA and for patiently listening to presentations about plasmids and genomes. A big thanks to everyone working in the anaerobic laboratory, especially **Ann-Louise Bergefur**, **Ulla Zimmerman** and **Gull Holmström**, for your cultivation work and for being a nice company during the many days of cultivating chicken caeca. A big thank you to **Marianne Persson** and **Boel Harbom** for taking care of things in the FoU-lab.

Tina for being an excellent exjobbare and running all those PCRs and PFGEs. I think your next "profession of today" should be to work closer to me again.

David, **Robert** and **Anna**, for being my roommates. Thanks to you I have learned about lucid dreaming and culturing of plants (and ocean creatures), shared the fascination for Jonas Karlsson and received last minute dinner

recipes. I have spent the majority of my eight years at SVA in your company and I have enjoyed it very much!

Work would be less productive without pauses spent with good colleagues. **Helena**, my bitter-sweet commuter-pal, work would be so boring without you. **Joann**, **Helena**, **Yvonne**, **Marie**, **Lena F**, and present and former "Bakt-kids": **Jocke**, **Sara F**, **Robert**, **Tina**, **Therese**, **Anders, Susanne, Sara Å, Tomas**, **Heidur** and **Sevinc** for all the laughs during coffee breaks, lunches and AWs.

MY COLLEAGUES AT SVA

To everyone at substrat (especially Jana and Åsa), who have made my cultivation media**. Stefan, Mattias** and **Oskar N** for sharing your knowledge about plasmids and other lab-oriented subjects at FoU-meetings. **Oskar K** for always being busy and exhausted from lack of sleep, but yet stopping to talk! **Anna** and **Åsa** for organizing PhD student "fika", and a crayfish party in November.

MY FRIENDS AND FAMILY

Linköpingskompisarna, **biologfamiljen**, **korridorsvännerna**, och övriga vänner, TACK för er fantastiska förmåga att distrahera mig från jobbet!

Ågrenarna, för go'-fikor, matlådor och barnvakt. Det är så mysigt att hänga med er!

Skarinarna och **Andénarna**, för era tålmodiga frågor om hur det går på jobbet egentligen. **Farfar**, nu är jag klar med den där vattenforskningen (som ju inte handlar om vatten) **Morfar**, jag vet att du är den som kommer lägga mest tid och energi på att läsa min avhandling. Tyvärr finns inte *Clostridium botulinum* som gosedjur, annars hade du fått ett. **Mamma** och **Pappa,** för allt vi har upplevt tillsammans och för att ni alltid trott att jag ska klara saker, även när jag själv är skeptisk. Jag önskar att jag kunde föra Östergötland närmare Uppland.

Elvin, för att du gett mig en anledning att åka hem i tid från jobbet för att hinna busa och kramas. Du är det bästa jag någonsin gjort.

Jocke, för att du är du och vi är vi. Och för att du gör mig så glad varje dag ☺

10 References

Agren, J., Sundstrom, A., Hafstrom, T. & Segerman, B. (2012). Gegenees: fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. *PLoS One,* 7 (6)**,** pp. e39107.

Aktories, K., Barmann, M., Ohishi, I., Tsuyama, S., Jakobs, K. H. & Habermann, E. (1986). Botulinum C2 toxin ADP-ribosylates actin. *Nature,* 322 (6077)**,** pp. 390-2.

Aktories, K., Braun, U., Rosener, S., Just, I. & Hall, A. (1989). The rho gene product expressed in *E. coli* is a substrate of botulinum ADP-ribosyltransferase C3. *Biochemical and Biophysical Research Communications,* 158 (1)**,** pp. 209-13.

Alikhan, N. F., Petty, N. K., Ben Zakour, N. L. & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics,* 12**,** pp. 402.

AniBioThreat. (2010). AniBioThreat. Available: http://www.anibiothreat.com/ [Accessed 2015-06-29].

Anniballi, F., Auricchio, B., Woudstra, C., Fach, P., Fiore, A., Skarin, H., Bano, L., Segerman, B., Knutsson, R. & De Medici, D. (2013). Multiplex real-time PCR for detecting and typing *Clostridium botulinum* group III organisms and their mosaic variants. *Biosecurity and bioterrorism : biodefense strategy, practice, and science,* 11 Suppl 1**,** pp. S207-14.

Arnon, S. S., Schechter, R., Inglesby, T. V., Henderson, D. A., Bartlett, J. G., Ascher, M. S., Eitzen, E., Fine, A. D., Hauer, J., Layton, M., Lillibridge, S., Osterholm, M. T., O'toole, T., Parker, G., Perl, T. M., Russell, P. K., Swerdlow, D. L. & Tonat, K. (2001). Botulinum toxin as a biological weapon: medical and public health management. *JAMA,* 285 (8)**,** pp. 1059-70.

Aureli, P., Fenicia, L., Pasolini, B., Gianfranceschi, M., Mccroskey, L. M. & Hatheway, C. L. (1986). Two cases of type E infant botulism caused by neurotoxigenic *Clostridium butyricum* in Italy. *The Journal of infectious diseases,* 154 (2)**,** pp. 207-11.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A.

V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A. & Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology : a journal of computational molecular cell biology,* 19 (5)**,** pp. 455-77.

Barash, J. R. & Arnon, S. S. (2014). A novel strain of *Clostridium botulinum* that produces type B and type H botulinum toxins. *The Journal of infectious diseases,* 209 (2)**,** pp. 183-91.

Bengtson, I. A. (1922). Preliminary note on a toxin producing anaerobe isolated from the larvae of *Lucilia caesar*. *Public Health Rep,* 37**,** pp. 164-170.

Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., Boutell, J. M., Bryant, J., Carter, R. J., Keira Cheetham, R., Cox, A. J., Ellis, D. J., Flatbush, M. R., Gormley, N. A., Humphray, S. J., Irving, L. J., Karbelashvili, M. S., Kirk, S. M., Li, H., Liu, X., Maisinger, K. S., Murray, L. J., Obradovic, B., Ost, T., Parkinson, M. L., Pratt, M. R., Rasolonjatovo, I. M., Reed, M. T., Rigatti, R., Rodighiero, C., Ross, M. T., Sabot, A., Sankar, S. V., Scally, A., Schroth, G. P., Smith, M. E., Smith, V. P., Spiridou, A., Torrance, P. E., Tzonev, S. S., Vermaas, E. H., Walter, K., Wu, X., Zhang, L., Alam, M. D., Anastasi, C., Aniebo, I. C., Bailey, D. M., Bancarz, I. R., Banerjee, S., Barbour, S. G., Baybayan, P. A., Benoit, V. A., Benson, K. F., Bevis, C., Black, P. J., Boodhun, A., Brennan, J. S., Bridgham, J. A., Brown, R. C., Brown, A. A., Buermann, D. H., Bundu, A. A., Burrows, J. C., Carter, N. P., Castillo, N., Chiara, E. C. M., Chang, S., Neil Cooley, R., Crake, N. R., Dada, O. O., Diakoumakos, K. D., Dominguez-Fernandez, B., Earnshaw, D. J., Egbujor, U. C., Elmore, D. W., Etchin, S. S., Ewan, M. R., Fedurco, M., Fraser, L. J., Fuentes Fajardo, K. V., Scott Furey, W., George, D., Gietzen, K. J., Goddard, C. P., Golda, G. S., Granieri, P. A., Green, D. E., Gustafson, D. L., Hansen, N. F., Harnish, K., Haudenschild, C. D., Heyer, N. I., Hims, M. M., Ho, J. T., Horgan, A. M., et al. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature,* 456 (7218)**,** pp. 53-9.

Binz, T., Blasi, J., Yamasaki, S., Baumeister, A., Link, E., Sudhof, T. C., Jahn, R. & Niemann, H. (1994). Proteolysis of SNAP-25 by types E and A botulinal neurotoxins. *The Journal of biological chemistry,* 269 (3)**,** pp. 1617-20.

Blasi, J., Chapman, E. R., Yamasaki, S., Binz, T., Niemann, H. & Jahn, R. (1993). Botulinum neurotoxin C1 blocks neurotransmitter release by means of cleaving HPC-1/syntaxin. *EMBO J,* 12 (12)**,** pp. 4821-8.

Blomqvist, G., Skarin, H., Lindberg, A., Båverud, V. & Engström, B. (2009). Surveillance for *Clostridium botulinum* type C or C/D in Swedish broilers. Proceedings of the 16th WVPA Congress, 2009 Marrakesh.

Bohnel, H., Schwagerick, B. & Gessler, F. (2001). Visceral botulism--a new form of bovine *Clostridium botulinum* toxication. *Journal of veterinary medicine. A, Physiology, pathology, clinical medicine,* 48 (6)**,** pp. 373-83.

Bruggemann, H., Baumer, S., Fricke, W. F., Wiezer, A., Liesegang, H., Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A. & Gottschalk, G.

(2003). The genome sequence of *Clostridium tetani*, the causative agent of tetanus disease. *Proceedings of the National Academy of Sciences of the United States of America,* 100 (3)**,** pp. 1316-21.

Burke, G. S. (1919). Notes on *Bacillus botulinus*. *Journal of bacteriology,* 4 (5)**,** pp. 555-570 1.

Carter, A. T., Paul, C. J., Mason, D. R., Twine, S. M., Alston, M. J., Logan, S. M., Austin, J. W. & Peck, M. W. (2009). Independent evolution of neurotoxin and flagellar genetic loci in proteolytic *Clostridium botulinum*. *BMC Genomics,* 10**,** pp. 115.

Carter, A. T. & Peck, M. W. (2015). Genomes, neurotoxins and biology of *Clostridium botulinum* Group I and Group II. *Research in microbiology,* 166 (4)**,** pp. 303-17.

Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M. A., Barrell, B. G. & Parkhill, J. (2005). ACT: the Artemis Comparison Tool. *Bioinformatics,* 21 (16)**,** pp. 3422-3.

Cato, E. P., Hash, D. E., Holdeman, L. V. & Moore, W. E. (1982). Electrophoretic study of *Clostridium* species. *Journal of clinical microbiology,* 15 (4)**,** pp. 688-702.

Centers for Disese Control and Prevention (CDC). (2012). National Botulism Surveillance System Overview. Available: http://www.cdc.gov/nczved/divisions/dfbmd/diseases/botulism/ [Accessed 2015-07-10].

Centers for Disease Control and Prevention (CDC). (2013). Standard Operating Procedure for PulseNet PFGE of *Clostridium botulinum*. Available: http://www.cdc.gov/pulsenet/PDF/c-botulinum-protocol-508c.pdf [Accessed 2015-07-13].

Chen, Y., Korkeala, H., Aarnikunnas, J. & Lindstrom, M. (2007). Sequencing the botulinum neurotoxin gene and related genes in *Clostridium botulinum* type E strains reveals orfx3 and a novel type E neurotoxin subtype. *Journal of bacteriology,* 189 (23)**,** pp. 8643-50.

Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A. J., Muller, W. E., Wetter, T. & Suhai, S. (2004). Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res,* 14 (6)**,** pp. 1147-59.

Collins, M. D. & East, A. K. (1998). Phylogeny and taxonomy of the food-borne pathogen *Clostridium botulinum* and its neurotoxins. *Journal of applied microbiology,* 84 (1)**,** pp. 5-17.

Critchley, E. M. (1991). A comparison of human and animal botulism: a review. *J R Soc Med,* 84 (5)**,** pp. 295-8.

Dahlenborg, M., Borch, E. & Radstrom, P. (2001). Development of a combined selection and enrichment PCR procedure for *Clostridium botulinum* Types B, E, and F and its use to determine prevalence in fecal samples from slaughtered pigs. *Applied and environmental microbiology,* 67 (10)**,** pp. 4781-8.

Darling, A. E., Mau, B. & Perna, N. T. (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One,* 5 (6)**,** pp. e11147.

Delcher, A. L., Harmon, D., Kasif, S., White, O. & Salzberg, S. L. (1999). Improved microbial gene identification with GLIMMER. *Nucleic acids research,* 27 (23)**,** pp. 4636-41.

Dinter, V. Z. & Kull, K. E. (1951). Ergänzende Untersuchungen über den Botulismus beim Nerz. *Nordisk veterinaermedicin,* 3**,** pp. 297-311.

Dong, M., Liu, H., Tepp, W. H., Johnson, E. A., Janz, R. & Chapman, E. R. (2008). Glycosylated SV2A and SV2B mediate the entry of botulinum neurotoxin E into neurons. *Molecular biology of the cell,* 19 (12)**,** pp. 5226-37.

Dong, M., Richards, D. A., Goodnough, M. C., Tepp, W. H., Johnson, E. A. & Chapman, E. R. (2003). Synaptotagmins I and II mediate entry of botulinum neurotoxin B into cells. *The journal of cell biology,* 162 (7)**,** pp. 1293-303.

Dong, M., Yeh, F., Tepp, W. H., Dean, C., Johnson, E. A., Janz, R. & Chapman, E. R. (2006). SV2 is the protein receptor for botulinum neurotoxin A. *Science,* 312 (5773)**,** pp. 592-6.

European Centre for Disease Prevention and Control (ECDC). (2014). Annual epidemiological report 2014: Food- and waterborne diseases and zoonoses. *ECDC SURVEILLANCE REPORT* [Online]. Available: http://ecdc.europa.eu/en/healthtopics/botulism/Pages/Epidemiological-data.aspx [Accessed 2015-07-10].

Eklund, M. W. & Poysky, F. T. (1974). Interconversion of type C and D strains of *Clostridium botulinum* by specific bacteriophages. *Applied microbiology,* 27 (1)**,** pp. 251-8.

Eklund, M. W., Poysky, F. T., Meyers, J. A. & Pelroy, G. A. (1974). Interspecies conversion of *Clostridium botulinum* type C to *Clostridium novyi* type A by bacteriophage. *Science,* 186 (4162)**,** pp. 456-8.

Eklund, M. W., Poysky, F. T., Oguma, K., Iida, H. & Inoue, K. (1987). Relationship of bacteriophages to toxin and hemagglutinin production by *Clostridium botulinum* types C and D and its significance in avian botulism outbreaks. *In:* EKLUND, M. W. & DOWELL, V. R. (eds.) *Avian botulism.* Springfield: Charles C Thomas.

Eklund, M. W., Poysky, F. T., Reed, S. M. & Smith, C. A. (1971). Bacteriophage and the toxigenicity of *Clostridium botulinum* type C. *Science,* 172 (3982)**,** pp. 480-2.

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M. & Et Al. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science,* 269 (5223)**,** pp. 496-512.

Foran, P., Lawrence, G. W., Shone, C. C., Foster, K. A. & Dolly, J. O. (1996). Botulinum neurotoxin C1 cleaves both syntaxin and SNAP-25 in intact and permeabilized chromaffin cells: correlation with its blockade of catecholamine release. *Biochemistry,* 35 (8)**,** pp. 2630-6.

Franz, D. R., Pitt, L. M., Clayton, M. A. & Hanes, M. A. (1993). Efficacy of prophylactic and therapeutic administration of antitoxin for inhalation botulism. *In:* DASGUPTA, B. R. (ed.) *Botulinum and Tetanus*

Neurotoxins: Neurotransmission and Biomedical Apsects. New York: Plenum Press.

Fu, Z., Chen, C., Barbieri, J. T., Kim, J. J. & Baldwin, M. R. (2009). Glycosylated SV2 and gangliosides as dual receptors for botulinum neurotoxin serotype F. *Biochemistry,* 48 (24)**,** pp. 5631-41.

Fujinaga, Y., Matsumura, T., Jin, Y., Takegahara, Y. & Sugawara, Y. (2009). A novel function of botulinum toxin-associated proteins: HA proteins disrupt intestinal epithelial barrier to increase toxin absorption. *Toxicon : official journal of the International Society on Toxinology,* 54 (5)**,** pp. 583-6.

GenBank. (2015). *Clostridium botulinum. National Center for Biotechnology Information (NCBI)* [Online]. Available: http://www.ncbi.nlm.nih.gov/genome/ [Accessed 2015-07-10].

Gill, D. M. (1982). Bacterial toxins: a table of lethal amounts. *Microbiological reviews,* 46 (1)**,** pp. 86-94.

Giltner, L. T. & Couch, J. F. (1930). Western Duck Sickness and Botulism. *Science,* 72 (1878)**,** pp. 660.

Gimenez, D. F. & Ciccarelli, A. S. (1970). Another type of *Clostridium botulinum. Zbl Bakt I abt Orig,* 215**,** pp. 212-220.

Gordon, D., Abajian, C. & Green, P. (1998). Consed: a graphical tool for sequence finishing. *Genome research,* 8 (3)**,** pp. 195-202.

Grate, J. W., Ozanich, R. M. J., Warner, M. G., Bruckner-Lea, C. J. & Marks, J. D. (2010). Advances in assays and analytical approaches for botulinum-toxin detection. *Trends in analytical chemistry,* 29**,** pp. 1137-1156.

Grenda, T., Kukier, E. & Kwiatek, K. (2014). Methods and difficulties in detection of *Clostridium botulinum* and its toxins. *Polish journal of veterinary sciences,* 17 (1)**,** pp. 195-205.

Gu, S., Rumpel, S., Zhou, J., Strotmeier, J., Bigalke, H., Perry, K., Shoemaker, C. B., Rummel, A. & Jin, R. (2012). Botulinum neurotoxin is shielded by NTNHA in an interlocked complex. *Science,* 335 (6071)**,** pp. 977-81.

Gunnison, J. B., Kummings, J. R. & Meyer, K. F. (1936). *Clostridium botulinum* type E. *Proceedings of the society for experimental biology and medicine.,* 35**,** pp. 278-280.

Hall, J. D., Mccroskey, L. M., Pincomb, B. J. & Hatheway, C. L. (1985). Isolation of an organism resembling *Clostridium barati* which produces type F botulinal toxin from an infant with botulism. *Journal of clinical microbiology,* 21 (4)**,** pp. 654-5.

Hariharan, H. & Mitchell, W. R. (1976). Observations on bacteriophages of *Clostridium botulinum* type C isolates from different sources and the role of certain phages in toxigenicity. *Applied and environmental microbiology,* 32 (1)**,** pp. 145-58.

Hatheway, C. L. (1990). Toxigenic clostridia. *Clinical micriobiology reviews,* 3 (1)**,** pp. 66-98.

Hauschild, A. H. W. (1989). *Clostridium botulinum. In:* DOYLE, M. P. (ed.) *Foodborne bacterial pathogens.* New York: Marcel Dekker, Inc.

Hauser, D., Gibert, M., Eklund, M. W., Boquet, P. & Popoff, M. R. (1993). Comparative analysis of C3 and botulinal neurotoxin genes and their

environment in *Clostridium botulinum* types C and D. *Journal of bacteriology,* 175 (22)**,** pp. 7260-8.

Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P. & Minton, N. P. (2007). The ClosTron: a universal gene knock-out system for the genus *Clostridium*. *Journal of microbiological methods,* 70 (3)**,** pp. 452-64.

Hedeland, M., Moura, H., Baverud, V., Woolfitt, A. R., Bondesson, U. & Barr, J. R. (2011). Confirmation of botulism in birds and cattle by the mouse bioassay and Endopep-MS. *Journal of medical microbiology,* 60 (Pt 9)**,** pp. 1299-305.

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.* Berlin Heidelberg: Springer-Verlag.

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 (3)**,** pp. 818-32.

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**,** pp. 66.

Hogg, R. A., White, V. J. & Smith, G. R. (1990). Suspected botulism in cattle associated with poultry litter. *The Veterinary record,* 126 (19)**,** pp. 476-9.

Holzer, V. E. (1962). Botulism from inhalation [in german]. *Med Klin,* 57**,** pp. 1735-1738.

Hooper, D. C. & Hirsch, M. S. (2014). Novel clostridium botulinum toxin and dual use research of concern issues. *The journal of infectious diseases,* 209 (2)**,** pp. 167.

Hunter, L. C., Miller, J. K. & Poxton, I. R. (1999). The association of *Clostridium botulinum* type C with equine grass sickness: a toxicoinfection? *Equine veterinary journal,* 31 (6)**,** pp. 492-9.

Huss, H. H. (1980). Distribution of *Clostridium botulinum*. *Applied and environmental microbiology,* 39 (4)**,** pp. 764-9.

Hyun, S. H. & Sakaguchi, G. (1989). Implication of coprophagy in pathogenesis of chicken botulism. *Nihon juigaku zasshi. The Japanese journal of veterinary science,* 51 (3)**,** pp. 582-6.

Illumina. (2015). Sequencing and array-based solutions for genetic research. Available: http://www.illumina.com/ [Accessed 2015-06-22].

Inoue, K. & Iida, H. (1970). Conversion of toxigenicity in *Clostridium botulinum* type C. *Japanese journal of microbiology,* 14**,** pp. 87-89.

Inui, K., Sagane, Y., Miyata, K., Miyashita, S., Suzuki, T., Shikamori, Y., Ohyama, T., Niwa, K. & Watanabe, T. (2012). Toxic and nontoxic components of botulinum neurotoxin complex are evolved from a common ancestral zinc protein. *Biochemical and biophysical research communications,* 419 (3)**,** pp. 500-4.

Ito, K., Seslar, D. & Ercern, W. (1967). The thermal and chlorine resistance of Clostridium botulinum types A, B, and E spores. *In:* INGRAM, M. & ROBERTS, T. (eds.) *Botulism.* London: Chapman & Hall.

Jensen, W. I. & Price, J. I. (1987). The global importance of type C botulism in wild birds. *In:* THOMAS, C. C. (ed.) *Avian Botulism: An International Perspective.* Springfield, Illinois.

Jin, Y., Takegahara, Y., Sugawara, Y., Matsumura, T. & Fujinaga, Y. (2009). Disruption of the epithelial barrier by botulinum haemagglutinin (HA) proteins - differences in cell tropism and the mechanism of action between HA proteins of types A or B, and HA proteins of type C. *Microbiology,* 155 (Pt 1)**,** pp. 35-45.

Johnson, A. L., Mcadams, S. C. & Whitlock, R. H. (2010). Type A botulism in horses in the United States: a review of the past ten years (1998-2008). *Journal of veterinary diagnostic investigation : official publication of the American Association of Veterinary Laboratory Diagnosticians, Inc,* 22 (2)**,** pp. 165-73.

Johnson, A. L., Sweeney, R. W., Mcadams, S. C. & Whitlock, R. H. (2012). Quantitative real-time PCR for detection of the neurotoxin gene of *Clostridium botulinum* type B in equine and bovine samples. *Veterinary journal,* 194 (1)**,** pp. 118-20.

Johnson, E. A. (1999). Clostridial toxins as therapeutic agents: benefits of nature's most toxic proteins. *Annual review of microbiology,* 53**,** pp. 551-75.

Kaji, R. (2011). New and emerging indications of botulinum toxin therapy. *Parkinsonism and related disorders,* 17 Suppl 1**,** pp. S25-7.

Kalb, S. R., Baudys, J., Smith, T. J., Smith, L. A. & Barr, J. R. (2014). Three enzymatically active neurotoxins of *Clostridium botulinum* strain Af84: BoNT/A2, /F4, and /F5. *Analytical chemistry,* 86 (7)**,** pp. 3254-62.

Kalb, S. R., Krilich, J. C., Dykes, J. K., Luquez, C., Maslanka, S. E. & Barr, J. R. (2015). Detection of Botulinum Toxins A, B, E, and F in Foods by Endopep-MS. *Journal of agricultural and food chemistry*.

Koren, S., Harhay, G. P., Smith, T. P., Bono, J. L., Harhay, D. M., Mcvey, S. D., Radune, D., Bergman, N. H. & Phillippy, A. M. (2013). Reducing assembly complexity of microbial genomes with single-molecule sequencing. *Genome biology,* 14 (9)**,** pp. R101.

Koriazova, L. K. & Montal, M. (2002). Translocation of botulinum neurotoxin light chain protease through the heavy chain channel. *Nature Structural Biology,* 10**,** pp. 13-18.

Kristjansson, M., Samore, M. H., Gerding, D. N., Degirolami, P. C., Bettin, K. M., Karchmer, A. W. & Arbeit, R. D. (1994). Comparison of restriction endonuclease analysis, ribotyping, and pulsed-field gel electrophoresis for molecular differentiation of *Clostridium difficile* strains. *Journal of clinical microbiology,* 32 (8)**,** pp. 1963-9.

Kubota, T., Shirakawa, S., Kozaki, S., Isogai, E., Isogai, H., Kimura, K. & Fujii, N. (1996). Mosaic type of the nontoxic-nonhemaggulutinin component gene in *Clostridium botulinum* type A strain isolated from infant botulism in Japan. *Biochemical and biophysical research communications,* 224 (3)**,** pp. 843-8.

Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C. & Salzberg, S. L. (2004). Versatile and open software for comparing large genomes. *Genome biology,* 5 (2)**,** pp. R12.

Lalli, G., Herreros, J., Osborne, S. L., Montecucco, C., Rossetto, O. & Schiavo, G. (1999). Functional characterisation of tetanus and botulinum neurotoxins binding domains. *Journal of cell science,* 112 (Pt 16)**,** pp. 2715-24.

Lan Chun, C., Kahn, C. I., Borchert, A. J., Byappanahalli, M. N., Whitman, R. L., Peller, J., Pier, C., Lin, G., Johnson, E. A. & Sadowsky, M. J. (2015). Prevalence of toxin-producing *Clostridium botulinum* associated with the macroalga Cladophora in three Great Lakes: growth and management. *The Science of the total environment,* 511**,** pp. 523-9.

Lee, W. H. & Riemann, H. (1970). Correlation of toxic and non-toxic strains of *Clostridium botulinum* by DNA composition and homology. *Journal of general micriobiology,* 60 (1)**,** pp. 117-23.

Leuchs, J. (1910). Beitrage zu kenntnis des toxins und antitoxins des *Bacillus botulinus. Ztschr Hyg Infekt,* 65**,** pp. 55-84.

Life Technologies. (2015). Ion Community. Available: http://ioncommunity.lifetechnologies.com/welcome [Accessed 2015-06-22].

Lindberg, A., Skarin, H., Knutsson, R., Blomqvist, G. & Baverud, V. (2010). Real-time PCR for *Clostridium botulinum* type C neurotoxin (BoNTC) gene, also covering a chimeric C/D sequence--application on outbreaks of botulism in poultry. *Vet Microbiol,* 146 (1-2)**,** pp. 118-23.

Lindstedt, B. A. (2005). Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. *Electrophoresis,* 26 (13)**,** pp. 2567-82.

Lindstrom, M. & Korkeala, H. (2006). Laboratory diagnostics of botulism. *Clinical microbiology reviews,* 19 (2)**,** pp. 298-314.

Lindstrom, M., Myllykoski, J., Sivela, S. & Korkeala, H. (2010). *Clostridium botulinum* in cattle and dairy products. *Critical reviews in food science and nutrition,* 50 (4)**,** pp. 281-304.

Lindstrom, M., Nevas, M., Kurki, J., Sauna-Aho, R., Latvala-Kiesila, A., Polonen, I. & Korkeala, H. (2004). Type C botulism due to toxic feed affecting 52,000 farmed foxes and minks in Finland. *Journal of clinical microbiology,* 42 (10)**,** pp. 4718-25.

Livesey, C. T., Sharpe, R. T. & Hogg, R. A. (2004). Recent association of cattle botulism with poultry litter. *The Veterinary record,* 154 (23)**,** pp. 734-5.

Mahillon, J. & Chandler, M. (1998). Insertion sequences. *Microbiology and molecular biology reviews: MMBR,* 62 (3)**,** pp. 725-74.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y. J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L., Jarvie, T. P., Jirage, K. B., Kim, J. B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., Mcdade, K. E., Mckenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F.,

Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F. & Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature,* 437 (7057)**,** pp. 376-80.

Marshall, K. M., Bradshaw, M. & Johnson, E. A. (2010). Conjugative botulinum neurotoxin-encoding plasmids in *Clostridium botulinum*. *PLoS One,* 5 (6)**,** pp. e11087.

Maslanka, S. E., Luquez, C., Dykes, J. K., Tepp, W. H., Pier, C. L., Pellett, S., Raphael, B. H., Kalb, S. R., Barr, J. R., Rao, A. & Johnson, E. A. (2015). A Novel Botulinum Toxin, Previously Reported as Serotype H, has a Hybrid Structure of Known Serotypes A and F that is Neutralized with Serotype A Antitoxin. *The Journal of infectious diseases*.

Mcclung, L. S. & Toabe, R. (1947). The Egg Yolk Plate Reaction for the Presumptive Diagnosis of *Clostridium sporogenes* and Certain Species of the Gangrene and Botulinum Groups. *Journal of bacteriology,* 53 (2)**,** pp. 139-47.

Mchugh, G. L. & Swartz, M. N. (1977). Elimination of plasmids from several bacterial species by novobiocin. *Antimicrobial agents and chemotherapy,* 12 (3)**,** pp. 423-6.

Meyer, K. F. & Gunnison, J. B. (1928). *Cl. Botulinum* type D. *Proceedings of the society for experimental biology and medicine,* 26**,** pp. 88-89.

Mitchell, W. R. & Rosendal, S. (1987). Type C botulism: The agent, host susceptibility, and predisposing factors. *In:* EKLUND, M. W. & DOWELL, V. R. (eds.) *Avian Botulism: An International Perspective.* Springfield, Illinois.

Miyazaki, S. & Sakaguchi, G. (1978). Experimental botulism in chickens: the cecum as the site of production and absorption of botulinum toxin. *Japanese journal of medical science & biology,* 31 (1)**,** pp. 1-15.

Moriishi, K., Koura, M., Abe, H. & Fujii, N. (1996a). Mosaic structure of neurotoxins produced from *Clostridium botulinum* types C and D organisms. *Biochimica et biophysica acta,* 1307**,** pp. 123-126.

Moriishi, K., Koura, M., Fujii, N., Fujinaga, Y., Inoue, K., Syuto, B. & Oguma, K. (1996b). Molecular cloning of the gene encoding the mosaic neurotoxin, composed of parts of botulinum neurotoxin types C1 and D, and PCR detection of this gene from *Clostridium botulinum* type C organisms. *Applied and environmental microbiology,* 62 (2)**,** pp. 662-7.

Myers, G. S., Rasko, D. A., Cheung, J. K., Ravel, J., Seshadri, R., Deboy, R. T., Ren, Q., Varga, J., Awad, M. M., Brinkac, L. M., Daugherty, S. C., Haft, D. H., Dodson, R. J., Madupu, R., Nelson, W. C., Rosovitz, M. J., Sullivan, S. A., Khouri, H., Dimitrov, G. I., Watkins, K. L., Mulligan, S., Benton, J., Radune, D., Fisher, D. J., Atkins, H. S., Hiscox, T., Jost, B. H., Billington, S. J., Songer, J. G., Mcclane, B. A., Titball, R. W., Rood, J. I., Melville, S. B. & Paulsen, I. T. (2006). Skewed genomic variability in strains of the toxigenic bacterial pathogen, *Clostridium perfringens*. *Genome research,* 16 (8)**,** pp. 1031-40.

Möller, V. & Scheibel, I. (1960). Preliminary report on the isolation of an apparently new type of *Cl. botulinum*. *Acta pathologica et microbiologica Scandinavica,* 48**,** pp. 80.

Nakamura, K., Kohda, T., Shibata, Y., Tsukamoto, K., Arimitsu, H., Hayashi, M., Mukamoto, M., Sasakawa, N. & Kozaki, S. (2012). Unique biological activity of botulinum D/C mosaic neurotoxin in murine species. *Infection and immunity,* 80 (8)**,** pp. 2886-93.

Nakamura, S., Kimura, I., Yamakawa, K. & Nishida, S. (1983). Taxonomic relationships among *Clostridium novyi* Types A and B, *Clostridium haemolyticum* and *Clostridium botulinum* type C. *Journal of general microbiology,* 129 (5)**,** pp. 1473-9.

Neimanis, A., Gavier-Widen, D., Leighton, F., Bollinger, T., Rocke, T. & Morner, T. (2007). An outbreak of type C botulism in herring gulls (*Larus argentatus*) in southeastern Sweden. *Journal of wildlife diseases,* 43 (3)**,** pp. 327-36.

Niwa, K., Koyama, K., Inoue, S., Suzuki, T., Hasegawa, K., Watanabe, T., Ikeda, T. & Ohyama, T. (2007). Role of nontoxic components of serotype D botulinum toxin complex in permeation through a Caco-2 cell monolayer, a model for intestinal epithelium. *FEMS immunology and medical microbiology,* 49 (3)**,** pp. 346-52.

Nolling, J., Breton, G., Omelchenko, M. V., Makarova, K. S., Zeng, Q., Gibson, R., Lee, H. M., Dubois, J., Qiu, D., Hitti, J., Wolf, Y. I., Tatusov, R. L., Sabathe, F., Doucette-Stamm, L., Soucaille, P., Daly, M. J., Bennett, G. N., Koonin, E. V. & Smith, D. R. (2001). Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *Journal of bacteriology,* 183 (16)**,** pp. 4823-38.

Notermans, S., Dufrenne, J. & Oosterom, J. (1981). Persistence of *Clostridium botulinum* type B on a cattle farm after an outbreak of botulism. *Applied and environmental microbiology,* 41 (1)**,** pp. 179-83.

Oakley, C. L. & Warrack, G. H. (1959). The soluble antigens of *Clostridium oedematiens* type D (*CI. haemolyticum*). *The Journal of pathology and bacteriology,* 78**,** pp. 543-51.

Octavia, S., Sara, J. & Lan, R. (2015). Characterization of a large novel phage-like plasmid in *Salmonella enterica* serovar *Typhimurium*. *FEMS microbiology letters,* 362 (8).

Oguma, K., Iida, H. & Inoue, K. (1975). Observations on nonconverting phage, c-n71, obtained from a nontoxigenic strain of *Clostridium botulinum* type C. *Japanese journal of microbiology,* 19 (3)**,** pp. 167-72.

Ortiz, N. E. & Smith, G. R. (1994). Landfill sites, botulism and gulls. *Epidemiology and infection,* 112 (2)**,** pp. 385-91.

Pacific Biosciences. (2015). PACBIO RS II Available: http://www.pacificbiosciences.com/ [Accessed 2015-06-22].

Peck, M. W., Stringer, S. C. & Carter, A. T. (2011). *Clostridium botulinum* in the post-genomic era. *Food microbiology,* 28 (2)**,** pp. 183-91.

Perez-Losada, M., Cabezas, P., Castro-Nallar, E. & Crandall, K. A. (2013). Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. *Infection, genetics and evolution : journal of molecular*

epidemiology and evolutionary genetics in infectious diseases, 16**,** pp. 38-53.

Popoff, M., Boquet, P., Gill, D. M. & Eklund, M. W. (1990). DNA sequence of exoenzyme C3, an ADP-ribosyltransferase encoded by *Clostridium botulinum* C and D phages. *Nucleic acids research,* 18 (5)**,** pp. 1291.

Raphael, B. H., Choudoir, M. J., Luquez, C., Fernandez, R. & Maslanka, S. E. (2010). Sequence diversity of genes encoding botulinum neurotoxin type F. *Applied and environmental microbiology,* 76 (14)**,** pp. 4805-12.

Roche. (2015). 454 sequencing. Available: http://www.454.com/ [Accessed 2015-06-22].

Rothberg, J. M., Hinz, W., Rearick, T. M., Schultz, J., Mileski, W., Davey, M., Leamon, J. H., Johnson, K., Milgrew, M. J., Edwards, M., Hoon, J., Simons, J. F., Marran, D., Myers, J. W., Davidson, J. F., Branting, A., Nobile, J. R., Puc, B. P., Light, D., Clark, T. A., Huber, M., Branciforte, J. T., Stoner, I. B., Cawley, S. E., Lyons, M., Fu, Y., Homer, N., Sedova, M., Miao, X., Reed, B., Sabina, J., Feierstein, E., Schorn, M., Alanjary, M., Dimalanta, E., Dressman, D., Kasinskas, R., Sokolsky, T., Fidanza, J. A., Namsaraev, E., Mckernan, K. J., Williams, A., Roth, G. T. & Bustillo, J. (2011). An integrated semiconductor device enabling non-optical genome sequencing. *Nature,* 475 (7356)**,** pp. 348-52.

Rummel, A., Hafner, K., Mahrhold, S., Darashchonak, N., Holt, M., Jahn, R., Beermann, S., Karnath, T., Bigalke, H. & Binz, T. (2009). Botulinum neurotoxins C, E and F bind gangliosides via a conserved binding site prior to stimulation-dependent uptake with botulinum neurotoxin F utilising the three isoforms of SV2 as second receptor. *Journal of neurochemistry,* 110 (6)**,** pp. 1942-54.

Rummel, A., Mahrhold, S., Bigalke, H. & Binz, T. (2004). The HCC-domain of botulinum neurotoxins A and B exhibits a singular ganglioside binding site displaying serotype specific carbohydrate interaction. *Molecular microbiology,* 51 (3)**,** pp. 631-43.

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 (48)**,** pp. 17472-7.

Sakaguchi, Y., Hayashi, T., Yamamoto, Y., Nakayama, K., Zhang, K., Ma, S., Arimitsu, H. & Oguma, K. (2009). Molecular analysis of an extrachromosomal element containing the C2 toxin gene discovered in *Clostridium botulinum* type C. *Journal of bacteriology,* 191 (10)**,** pp. 3282-91.

Sasaki, Y., Takikawa, N., Kojima, A., Norimatsu, M., Suzuki, S. & Tamura, Y. (2001). Phylogenetic positions of *Clostridium novyi* and *Clostridium haemolyticum* based on 16S rDNA sequences. *International journal of systematic and evolutionary microbiology,* 51 (Pt 3)**,** pp. 901-4.

Schiavo, G., Benfenati, F., Poulain, B., Rossetto, O., Polverino De Laureto, P., Dasgupta, B. R. & Montecucco, C. (1992). Tetanus and botulinum-B

neurotoxins block neurotransmitter release by proteolytic cleavage of synaptobrevin. *Nature,* 359 (6398)**,** pp. 832-5.

Schiavo, G., Malizio, C., Trimble, W. S., Polverino De Laureto, P., Milan, G., Sugiyama, H., Johnson, E. A. & Montecucco, C. (1994). Botulinum G neurotoxin cleaves VAMP/synaptobrevin at a single Ala-Ala peptide bond. *The Journal of biological chemistry,* 269 (32)**,** pp. 20213-6.

Schiavo, G., Matteoli, M. & Montecucco, C. (2000). Neurotoxins affecting neuroexocytosis. *Physiological reviews,* 80 (2)**,** pp. 717-66.

Schiavo, G., Rossetto, O., Catsicas, S., Polverino De Laureto, P., Dasgupta, B. R., Benfenati, F. & Montecucco, C. (1993a). Identification of the nerve terminal targets of botulinum neurotoxin serotypes A, D, and E. *The Journal of biological chemistry,* 268 (32)**,** pp. 23784-7.

Schiavo, G., Shone, C. C., Rossetto, O., Alexander, F. C. & Montecucco, C. (1993b). Botulinum neurotoxin serotype F is a zinc endopeptidase specific for VAMP/synaptobrevin. *The Journal of biological chemistry,* 268 (16)**,** pp. 11516-9.

Sebaihia, M., Peck, M. W., Minton, N. P., Thomson, N. R., Holden, M. T., Mitchell, W. J., Carter, A. T., Bentley, S. D., Mason, D. R., Crossman, L., Paul, C. J., Ivens, A., Wells-Bennik, M. H., Davis, I. J., Cerdeno-Tarraga, A. M., Churcher, C., Quail, M. A., Chillingworth, T., Feltwell, T., Fraser, A., Goodhead, I., Hance, Z., Jagels, K., Larke, N., Maddison, M., Moule, S., Mungall, K., Norbertczak, H., Rabbinowitsch, E., Sanders, M., Simmonds, M., White, B., Whithead, S. & Parkhill, J. (2007). Genome sequence of a proteolytic (Group I) *Clostridium botulinum* strain Hall A and comparative analysis of the clostridial genomes. *Genome research,* 17 (7)**,** pp. 1082-92.

Sebaihia, M., Wren, B. W., Mullany, P., Fairweather, N. F., Minton, N., Stabler, R., Thomson, N. R., Roberts, A. P., Cerdeno-Tarraga, A. M., Wang, H., Holden, M. T., Wright, A., Churcher, C., Quail, M. A., Baker, S., Bason, N., Brooks, K., Chillingworth, T., Cronin, A., Davis, P., Dowd, L., Fraser, A., Feltwell, T., Hance, Z., Holroyd, S., Jagels, K., Moule, S., Mungall, K., Price, C., Rabbinowitsch, E., Sharp, S., Simmonds, M., Stevens, K., Unwin, L., Whithead, S., Dupuy, B., Dougan, G., Barrell, B. & Parkhill, J. (2006). The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nature genetics,* 38 (7)**,** pp. 779-86.

Segner, W. P. & Schmidt, C. F. (1971). Heat resistance of spores of marine and terrestrial strains of *Clostridium botulinum* type C. *Applied microbiology,* 22 (6)**,** pp. 1030-3.

Segner, W. P., Schmidt, C. F. & Boltz, J. K. (1971). Enrichment, isolation, and cultural characteristics of marine strains of *Clostridium botulinum* type C. *Applied microbiology,* 22 (6)**,** pp. 1017-24.

Seyboldt, C., Discher, S., Jordan, E., Neubauer, H., Jensen, K. C., Campe, A., Kreienbrock, L., Scheu, T., Wichern, A., Gundling, F., Doduc, P., Fohler, S., Abdulmawjood, A., Klein, G. & Hoedemaker, M. (2015). Occurrence of *Clostridium botulinum* neurotoxin in chronic disease of dairy cows. *Veterinary microbiology,* 177 (3-4)**,** pp. 398-402.

Shin, N. R., Byun, S. H., Chun, J. H., Shin, J. H., Kim, Y. J., Kim, J. H., Rhie, G. E., Chung, H. M., Mo, I. P. & Yoo, C. K. (2010). An outbreak of type C botulism in waterbirds: Incheon, Korea. *Journal of wildlife diseases,* 46 (3)**,** pp. 912-7.

Siegel, L. S. (1993). Destruction of botulinum toxin in food and water. *In:* HAUSCHILD, A. H. & DODDS, K. L. (eds.) *Clostridium botulinum: Ecology and Control in Foods.* New York: Marcel Dekker Inc.

Skarin, H., Lindgren, Y. & Jansson, D. S. (2015). Investigations into an Outbreak of Botulism Caused by *Clostridium botulinum* Type C/D in Laying Hens. *Avian Diseases,* 59**,** pp. 335-340.

Skarin, H., Tevell Aberg, A., Woudstra, C., Hansen, T., Lofstrom, C., Koene, M., Bano, L., Hedeland, M., Anniballi, F., De Medici, D. & Olsson Engvall, E. (2013). The workshop on animal botulism in Europe. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science,* 11 Suppl 1**,** pp. S183-90.

Smith, L. D. S. & Hobbs, G. (1974). Genus III Clostridium Prazmowski 1880, 23. . *In:* BUCHANAN, R. E. & GIBBONS, N. E. (eds.) *Bergey's Manual of Determinative Bacteriology.* 8th edition ed. Baltimore: William & Wilkins.

Smith, T. J., Lou, J., Geren, I. N., Forsyth, C. M., Tsai, R., Laporte, S. L., Tepp, W. H., Bradshaw, M., Johnson, E. A., Smith, L. A. & Marks, J. D. (2005). Sequence variation within botulinum neurotoxin serotypes impacts antibody binding and neutralization. *Infection and immunity,* 73 (9)**,** pp. 5450-7.

Sonnabend, O., Sonnabend, W., Heinzle, R., Sigrist, T., Dirnhofer, R. & Krech, U. (1981). Isolation of *Clostridium botulinum* type G and identification of type G botulinal toxin in humans: report of five sudden unexpected deaths. *The Journal of infectious diseases,* 143 (1)**,** pp. 22-7.

Stackebrandt, E., Kramer, I., Swiderski, J. & Hippe, H. (1999). Phylogenetic basis for a taxonomic dissection of the genus Clostridium. *FEMS immunology and medical microbiology,* 24 (3)**,** pp. 253-8.

Stringer, S. C., Carter, A. T., Webb, M. D., Wachnicka, E., Crossman, L. C., Sebaihia, M. & Peck, M. W. (2013). Genomic and physiological variability within Group II (non-proteolytic) *Clostridium botulinum*. *BMC Genomics,* 14**,** pp. 333.

Strotmeier, J., Lee, K., Volker, A. K., Mahrhold, S., Zong, Y., Zeiser, J., Zhou, J., Pich, A., Bigalke, H., Binz, T., Rummel, A. & Jin, R. (2010). Botulinum neurotoxin serotype D attacks neurons via two carbohydrate-binding sites in a ganglioside-dependent manner. *The Biochemical journal,* 431 (2)**,** pp. 207-16.

Sugawara, Y., Matsumura, T., Takegahara, Y., Jin, Y., Tsukasaki, Y., Takeichi, M. & Fujinaga, Y. (2010). Botulinum hemagglutinin disrupts the intercellular epithelial barrier by directly binding E-cadherin. *The Journal of cell biology,* 189 (4)**,** pp. 691-700.

Sunagawa, H., Ohyama, T., Watanabe, T. & Inoue, K. (1992). The complete amino acid sequence of the *Clostridium botulinum* type D neurotoxin, deduced

by nucleotide sequence analysis of the encoding phage d-16 phi genome. *J Vet Med Sci,* 54 (5)**,** pp. 905-13.

Swerczek, T. W. (1980). Toxicoinfectious botulism in foals and adult horses. *J Am Vet Med Assoc,* 176 (3)**,** pp. 217-20.

Szabo, E. A., Pemberton, J. M. & Desmarchelier, P. M. (1993). Detection of the genes encoding botulinum neurotoxin types A to E by the polymerase chain reaction. *Applied and environmental microbiology,* 59 (9)**,** pp. 3011-20.

Takeda, M., Tsukamoto, K., Kohda, T., Matsui, M., Mukamoto, M. & Kozaki, S. (2005). Characterization of the neurotoxin produced by isolates associated with avian botulism. *Avian Diseases,* 49 (3)**,** pp. 376-81.

Tevell Aberg, A., Bjornstad, K. & Hedeland, M. (2013). Mass spectrometric detection of protein-based toxins. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science,* 11 Suppl 1**,** pp. S215-26.

Tsukamoto, K., Kohda, T., Mukamoto, M., Takeuchi, K., Ihara, H., Saito, M. & Kozaki, S. (2005). Binding of *Clostridium botulinum* type C and D neurotoxins to ganglioside and phospholipid. Novel insights into the receptor for clostridial neurotoxins. *The Journal of biological chemistry,* 280 (42)**,** pp. 35164-71.

Van Ermengem, E. (1979). Classics in infectious diseases. A new anaerobic bacillus and its relation to botulism (translation of original 1897 article). *Reviews of infectious diseases,* 1**,** pp. 701-719.

(2013). Botulism: Fact sheet N°270. Available: http://www.who.int/mediacentre/factsheets/fs270/en/ [Accessed 2015-07-10].

Vidal, D., Anza, I., Taggart, M. A., Perez-Ramirez, E., Crespo, E., Hofle, U. & Mateo, R. (2013). Environmental factors influencing the prevalence of a *Clostridium botulinum* type C/D mosaic strain in nonpermanent Mediterranean wetlands. *Applied and environmental microbiology,* 79 (14)**,** pp. 4264-71.

Willjes, G., Mahrhold, S., Strotmeier, J., Eichner, T., Rummel, A. & Binz, T. (2013). Botulinum neurotoxin G binds synaptotagmin-II in a mode similar to that of serotype B: tyrosine 1186 and lysine 1191 cause its lower affinity. *Biochemistry,* 52 (22)**,** pp. 3930-8.

Windingstad, R.M. and Laitman, C.J. (1988). Ducks get sick, too! Available: http://www.nwhc.usgs.gov/publications/pamphlets/ducks_get_sick_pamphlet.pdf [Accessed 2015-05-27].

Winslow, C.-E., Broadhurst, J., Buchanan, R., Krumwiede, C., Rogers, L. & Smith, G. (1917). The families and genera of the bacteria- preliminary report of the committee of the society of American bacteriologists on characterization and classification of bacterial types. *Journal of bacteriology,* 2**,** pp. 505-566.

Woudstra, C., Le Marechal, C., Souillard, R., Bayon-Auboyer, M. H., Anniballi, F., Auricchio, B., De Medici, D., Bano, L., Koene, M., Sansonetti, M. H., Desoutter, D., Hansbauer, E. M., Dorner, M. B., Dorner, B. G. & Fach, P. (2015). Molecular gene profiling of *Clostridium botulinum* group III and its detection in naturally contaminated samples originating from various

European countries. *Applied and environmental microbiology,* 81 (7)**,** pp. 2495-505.

Woudstra, C., Skarin, H., Anniballi, F., Fenicia, L., Bano, L., Drigo, I., Koene, M., Bayon-Auboyer, M. H., Buffereau, J. P., De Medici, D. & Fach, P. (2012). Neurotoxin Gene Profiling of *Clostridium botulinum* Types C and D Native to Different Countries within Europe. *Applied and environmental microbiology,* 78 (9)**,** pp. 3120-7.

Vuylsteke, M., Peleman, J. D. & Van Eijk, M. J. (2007). AFLP technology for DNA fingerprinting. *Nature protocols,* 2 (6)**,** pp. 1387-98.

Yule, A. M., Barker, I. K., Austin, J. W. & Moccia, R. D. (2006). Toxicity of *Clostridium botulinum* type E neurotoxin to Great Lakes fish: implications for avian botulism. *Journal of wildlife diseases,* 42 (3)**,** pp. 479-93.

Zhong, L., Cheng, Q., Tian, X., Zhao, L. & Qin, Z. (2010). Characterization of the replication, transfer, and plasmid/lytic phage cycle of the *Streptomyces* plasmid-phage pZL12. *Journal of bacteriology,* 192 (14)**,** pp. 3747-54.