Spring 4-27-2022

Genomic Analysis of Metabolic Differences Found in *Clostridium perfringens* That Cause Necrotic Enteritis in Poultry

Connor Aylor
University of Nebraska-Lincoln, connor.aylor@huskers.unl.edu

GENOMIC ANALYSIS OF METABOLIC DIFFERENCES FOUND IN *CLOSTRIDIUM PERFRINGENS* THAT CAUSE NECROTIC ENTERITIS IN POULTRY

by

Connor M. Aylor

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Veterinary Science

Under the Supervision of Professors Greg Somerville and Etsuko Moriyama

Lincoln, Nebraska

April, 2022

GENOMIC ANALYSIS OF METABOLIC DIFFERENCES FOUND IN *CLOSTRIDIUM PERFRINGENS* THAT CAUSE NECROTIC ENTERITIS IN POULTRY

Connor M. Aylor, M.S.

University of Nebraska, 2022

Advisors: Greg Somerville and Etsuko Moriyama

Clostridium perfringens is a common member of gut microbiota in healthy animals, but can also be an important pathogen in human and veterinary medicine. It produces several protein toxins that contribute to both histotoxic and enteric diseases in animals. Necrotic enteritis in poultry has been associated with the NetB toxin of *C. perfringens*; however, this toxin alone is insufficient to cause disease in infected chickens. While considerable research has focused on the presence of toxins and virulence factors, little has been done to assess the function of metabolic factors on the ability of the bacteria to cause disease. In this study, the metabolic differences are examined using genomic sequence analysis between genomes of stains of *C. perfringens* that are associated with necrotic enteritis. Several metabolic pathways are examined, which show different metabolic genes across *C. perfringens* lineages.

AKNOWLEDGEMENTS

First, I would like to thank my major advisor Dr. Greg Somerville and co-advisor Dr. Etsuko Moriyama, for supporting me through an unconventional time period. Beyond their scientific expertise they helped me make the most of suboptimal circumstances and pushed me to keep driving toward my goals.

I would also like to thank the other members of my committee. Dr. J. Dustin Loy, for being willing to sit down and talk with me as I started this research, and providing me direction and resources. Dr. Rodrigo Franco Cruz, for making difficult concepts easy, and improving my literature search skills.

In addition, I would like to thank Dr. Raul Barletta for his part in the review portion of this thesis and for encouraging me to study areas of personal interest and think outside the box.

Finally, I would like to thank my family. My parents, for their constant support of my pursuits. Devin, for telling me it doesn't matter. Andy, Xav, Ayven, and Memi, for their unconditional love and support. Evan, for constantly showing me what strength looks like. Last and least, Aliece, for knowing what she wants and helping me know what I want.

TABLE OF CONTENTS

ABSTRACT....ii
AKNOWLEDGEMENTS....iii
CHAPTER I INTRODUCTION....1
1.1 Toxins Produced by *Clostridium perfringens*....1
1.1.1 Typing Toxins....3
1.1.2 Non-Typing Toxins....9
1.2 Non-Toxigenic Virulence Determinants....11
1.2.1 Agr-like Quorum Sensing....11
1.2.2 VirR/VirS System....11
1.2.3 Adhesins....12
1.3 *C. perfringens* Genomes....12
1.4 Current Problems in *C. perfringens* Genome and Metabolic Pathway Research....13
1.5 Objectives of this Study....14
Chapter II MATERIALS AND METHODS....15
2.1 Strain Selection....15
2.2 Phylogenetic Analysis....16
2.3 Visualization of Ortholog Distribution among the Genomes....17
2.4 Comparative Proteomic Analysis....17
2.5 Metabolic Pathway Analysis....18
Chapter III RESULTS AND DISCUSSION....19
3.1 Selection of *C. perfringens* Strains and Genomes....19
3.2 Phylogenetic Relationships of Selected *C. perfringens* genomes....19
3.3 Distribution of Orthologs in *C. perfringens* genomes....20
3.4 Comparison of Protein Families among *C. perfringens* Genomes....22
3.5 Metabolic Pathway Analysis among *C. perfringens* Genomes....23
3.5.1 Carbohydrate Biosynthesis Pathways....24
3.5.2 Ornithine Cycle....27
3.5.3 Phospholipase C Pathway....27
3.5.4 Carboxylate Degradation Pathways....28
3.5.5 Amino Acid Biosynthesis Pathways....30
3.5.6 Amine and Polyamine Degradation Pathways....30
3.5.7 Alcohol Degradation Pathways....32

3.5.8 Phosphopantothenate Biosynthesis 1 Pathways 33
3.6 Discussion 34
Chapter IV CONCLUSIONS AND FUTURE STUDIES 36
References 37
APPENDIX A: Genome Statistics 47
APPENDIX B: Phylogenetic reconstruction 51

LIST OF MULTIMEDIA CONTENT

CHAPTER 1

Table 1: Toxinotypes of *C. perfringens*.....2

Table 2: Toxins and disease of *C. perfringens* toxinotypes.....2

CHAPTER 2

Table 3: Genomic sequences used in this study and associated information.....16

CHAPTER 3

Figure 1: Phylogenetic reconstruction of selected *C. perfringens* strain genomes.....19

Figure 2: Ortholog distribution among selected C. perfringens genomes 21

Figure 3: Metabolic pathway distribution among the 19 *C. perfringens* genomes.....23

Figure 4: UDP-α-D-glucose biosynthesis pathway.....25

Figure 5: Frame-shifting insertion found in the phospho-sugar mutase gene in TAM-NE40.....25

Figure 6: dTDP-β-L-rhamnose biosynthesis pathway.....26

Figure 7: Ornithine cycle.....27

Figure 8: Nonsense mutation found in the *cpa* gene in UDE-95-1372.....28

Figure 9: β-D-glucuronide and D-glucuronate degradation pathway.....29

Figure 10: D-fructuronate degradation pathway.....29

Figure 11: Alignment of the aspartate kinase gene in TAM-NE40.....30

Figure 12: Creatinine degradation pathway.....31

Figure 13: Truncation of creatininase in Strain 37 and Strain 67.....32

Figure 14: L-lactaldehyde degradation pathway.....33

Figure 15: Frame-shifting mutation and truncation of 2-dehydropantoate 2-reductase in TAM-NE40.....34

APPENDIX A

Table A1: Isolation data of selected strains 47

Table A2: Sequencing data of selected strains 48

Table A3: Genomic data of selected strains 49

Table A4: Annotation data of selected strains ……..……50

APPENDIX B

Figure B1: Phylogenetic reconstruction using *C. ventriculi* as outgroup……53

CHAPTER I INTRODUCTION

Clostridium perfringens is a Gram-positive obligate anaerobe, known for its ability to sporulate and produce many toxins (Kiu and Hall 2018). *C. perfringens* is found ubiquitously in soil and water and is a common member of gut microbiota in healthy animals (Rood, Adams, et al. 2018). Due to its presence in gut flora, *C. perfringens* can induce enteric diseases when the health of the host changes. In addition to enteric diseases, *C. perfringens* causes several important histotoxic diseases, including gas gangrene. Importantly, the diversity of diseases and their severity are correlated with the large number of bacterial toxins found in the genome and on plasmids in *C. perfringens* strains (Li, Adams, et al. 2013).

1.1 Toxins Produced by *Clostridium perfringens*

Depending on its genomic and plasmid content, *C. perfringens* synthesizes more than 20 different toxins (Kiu and Hall 2018). This diversity in toxin content led to the commonly used toxinotyping scheme that was established in the 1960s. It was centered on the ability of strains to produce Alpha toxin (CPA), Beta toxin (CPB), Epsilon toxin (ETX) and Iota toxin (ITX) (Li, Adams, et al. 2013). As new toxins were identified and the diseases they cause were described, these were incorporated into this toxinotyping scheme. These additional toxins include enterotoxin (CPE) and necrotic enteritis beta-like toxin (NetB). This expanded repertoire of toxins has led to the identification of seven toxinotypes, categorized as A through G (Rood, Adams, et al. 2018). The seven toxinotypes and the major toxins associated with these toxinotypes are summarized in Tables 1 and 2.

Table 1: Toxinotypes of *C. perfringens*[a]

TOXINOTYPE	CPA	CPB	ETX	ITX	CPE	NETB
A	+	-	-	-	**-**	**-**
B	+	+	+	-	**-**	**-**
C	+	+	-	-	**+/-**	**-**
D	+	-	+	-	**+/-**	**-**
E	+	-	-	+	**+/-**	**-**
F	**+**	**-**	**-**	**-**	**+**	**-**
G	**+**	**-**	**-**	**-**	**-**	**+**

[a]The original toxinotyping system (A-E; shown with the grey background) as well as the extension (F and G) suggested by Rood et al. (2018) and the toxins produced by each toxinotype are listed.

Table 2: Toxins and disease of *C. perfringens* toxinotypes

Type	Toxins	Disease (Host)
A	CPA	Gas Gangrene (Most Mammals), Yellow Lamb Disease (Sheep)
B	CPA, CPB, ETX	Lamb Dysentery (Sheep)
C	CPA, CPB	Hemorrhagic/necrotizing enteritis (Many mammals)
D	CPA, ETX	Enterotoxemia (Sheep, goats and Cattle)
E	CPA, ITX	Hemorrhagic Gastroenteritis (Bovine)
F	CPA, CPE	Food Poisoning (Human)
G	CPA, NETB	Necrotic Enteritis (Poultry)

1.1.1 Typing Toxins

i. Alpha Toxin (Phospholipase C)

Alpha toxin (CPA) is produced by all toxinotypes of *C. perfringens*. CPA is a 370 amino acid zinc metallophospholipase that has two catalytic activities: phospholipase C and sphingomyelinase (Sakurai, Nagahama and Oda 2004). In neutrophils, Interleukin-8 (IL-8) release is induced through tyrosine kinase A (TrkA)-mediated activation of extracellular regulated kinase 1/2 (ERK1/2) and nuclear factor kappa B (NF-κB). CPA also induces the p38 mitogen-activated protein kinase (MAPK) pathways (Oda, Shiihara, et al. 2012). As CPA concentrates in cell membranes, these membranes become unstable due to the cleavage of phosphatidylcholine head groups, causing leakage of anions from affected epithelial cells (Rehman, et al. 2006).

The CPA protein contains two domains important for activity. The N-terminal domain contains the active site responsible for enzymatic activity, while the C-terminal domain provides the ability of the toxin to bind to the lipid rafts in the presence of the ganglioside GM1a/TrkA complex in host cells (Oda, Terao, et al. 2015). While both domains are required for cytolysis of host cells, only the C-terminal domain elicits a protective immune response. For this reason, the C-terminal domain of CPA has been incorporated into vaccines, such as the Boehringer Ingelheim Alpha vaccines (Goossens, Verherstraeten, et al. 2016).

CPA is essential for many of the diseases caused by *C. perfringens*; however, full bacterial virulence requires other toxins. This is evident from CPA's synergistic effects with Perfringolysin O (PFO) that enhances disease progression of gas gangrene (Awad, et al. 2001) as well as intestinal disease (Goossens, Valgaeren, et al. 2017). Injection of

CPA impairs the ability of neutrophils to differentiate, leading to a decrease in mature neutrophils both at the site of infection and in peripheral circulation reducing the ability of the immune system to contain *C. perfringens* infections (Takehara, et al. 2016).

C. perfringens toxinotype A (Table 1) is hypothesized to be involved with several gastrointestinal diseases in mammals. In contrast to other *C. perfringens* diseases that are associated with specific toxins, few diseases are directly attributable to alpha-toxin alone, leading to the hypothesis that alpha-toxin's primary function is synergistic (Uzal and Songer 2008). The major exception is yellow lamb disease, an extremely rare acute enterotoxemia seen in sheep and goats, which is associated with high levels of CPA (Giannitti, et al. 2014). This disease is characterized by non-specific symptoms of severe anemia, hemoglobinuria, and icterus. Despite the correlation between yellow lamb disease and the presence of high levels of CPA, causation is difficult to establish due to the presence of *C. perfringens* and CPA in normal gut microbiota of healthy animals (Uzal and Songer 2008).

ii. Beta Toxin

Beta toxin (CPB) is a 336 amino acid, plasmid borne, pore-forming toxin produced by toxinotype B and C strains of *C. perfringens*. Strains of *C. perfringens* that produce CPB without producing ETX are toxinotype C, while those producing both the CPB and ETX toxins are toxinotype B (Table 1). The CPB protein has 28% amino acid sequence similarity to *Staphylococcus aureus* alpha-toxin. They appear to function similarly through the creation of unregulated ion channels in cell membranes (Hunter, et al. 1993). CPB targets intestinal endothelial cells (Miclard, et al. 2009); however, the identity of the CPB receptor is unknown. That being said, CPB does interact with the purinergic receptor $P2X_7$ to induce cell death. Death of endothelial cells results in

vascular necrosis reducing blood flow and creating a hypoxic environment (Miclard, et al. 2009). The ability to induce cell death has established CPB as an important mediator of necrotic enteritis disease progression in several animal species. It should be noted that CPB fulfills the molecular Koch's postulates for necrotic enteritis, as CPB alone induces necrotic enteritis in animals. In addition, neutralization of CPB prevents the disease (Sayeed, et al. 2007).

C. perfringens toxinotype C, which produces only CPB but not ETX, causes intestinal infections in both humans and animals, most prominently in neonatal lambs. Beta toxin is susceptible to trypsin degradation; hence the lack of trypsin in neonates is associated with increased beta-toxin cytotoxic effects (Sayeed, et al. 2007). The disease is characterized by intestinal necrotic lesions most likely caused by beta-toxin endothelial damage or by direct damage to epithelium (Miclard, et al. 2009). Occasionally, infected sheep show few or no symptoms before sudden death in adults, a condition known as "struck" (McEwen and Roberts 1931).

iii. Epsilon Toxin

Epsilon toxin (ETX) is another pore-forming toxin produced by toxinotypes B and D (Rood, Adams, et al. 2018) (Table 1). It affects a wide range of cell types. For example, it accumulates in mouse kidney cells (Tamai, et al. 2003). It also affects lung epithelial cells (Dorca-Arevalo, et al. 2020). In addition, it is able to cross the blood brain barrier and affect an array of cells in the brain (Freedman, McClane and Uzal 2016), leading to speculation that it functions in inducing multiple sclerosis (Wagley, et al. 2019). ETX is a member of the aerolysin family of toxins that is secreted as a protoxin and activated by cleavage of both N- and C-terminal peptides. ETX is the third most potent toxin of the *Clostridium* spp., with an LD_{50} of ~70 ng/Kg (Minami, et al. 1997)

and a lethal dose below only the neurotoxins produced by *Clostridium tetani* and *Clostridium botulinum* (Popoff 2011). Speculation about the receptor for ETX has focused on the hepatitis A virus cellular receptor 1 (HAVCR1) (Ivie and McClain 2012); however, pores have been shown to form in artificial liposomes lacking the HAVCR1 as well (Nagahama, Hara, et al. 2006). After cell binding, ETX oligomerizes, a process requiring Caveolins 1 and 2 (Fennessey, et al. 2012). Epsilon toxin then inserts itself into the membrane and allows the unregulated passage of ions across the cell membrane, leading to cell death (Nestorovich, Karginov and Bezrukov 2010). Unlike CPB, which is inactivated by trypsin, ETX is activated by trypsin. This results in *C. perfringens* toxinotype B (containing both toxins) having a two-pronged attack when causing disease. Animals with low trypsin production often succumb to necrotic lesions of the intestine caused by CPB (as seen in toxinotype C infections), while at higher trypsin concentrations, ETX may play a lead role in pathogenesis similar to toxinotype D infections (Li, Adams, et al. 2013).

iv. Iota Toxin

Iota toxin (ITX) is a plasmid encoded A-B toxin (subunits are Ia and Ib) that is expressed by toxinotype E strains. ITX is closely related to the A-B toxins produced by *Clostridioides difficile* and *Clostridium spiroforme* (Redondo, et al. 2017). As with most A-B toxins, Ia is the active enzymatic component, while Ib is the binding component. Ib is believed to bind to Lipolysis-simulated lipoprotein receptor (LSR), similar to the A-B toxins from *C. difficile* (Papatheodorou, et al. 2011). Oligomerization of seven Ib subunits and binding of the Ia subunit facilitates endocytosis and translocation of Ia into the cytoplasm (Nagahama, Yamaguchi, et al. 2004). Translocation of Ia from the endosome requires acidification and a membrane potential gradient (Gibert, Marvaud, et

al. 2007). The Ia unit ADP-ribosylates actin (Vandekerckhove, et al. 1987), leading to the disruption of cellular structure and the release of cytochrome C from mitochondria (Nagahama, Ohkubo, et al. 2011). ITX is closely associated with hemorrhagic enteritis due to increasing intestinal permeability and destruction of the intestinal epithelium (Redondo, et al. 2017).

C. perfringens toxinotype E is a relatively uncommon toxinotype that produces ITX in addition to CPA. Toxinotype E is most commonly associated with hemorrhagic enteritis in neonatal calves (Diancourt, et al. 2019), goats (Kim, et al. 2013), ostriches (Keokilwe, et al. 2015), and possibly rabbits. Although, *C. spiroforme* may be the cause of the disease in rabbits (LaMont, Sonnenblick and Rothman 1979). While bovine enterotoxemia caused by *C. perfringens* toxinotype E is generally considered a disease of calves, in an Argentinian outbreak of enterotoxemia, unique strains that are able to cause disease in adult cattle were identified (Redondo, et al. 2013).

v. Enterotoxin

Enterotoxin (CPE) is a 319 amino acid protein and another pore-forming toxin in the aerolysin family (Briggs, et al. 2011). CPE is responsible for food poisoning in humans and is associated with enteric diseases in non-human animals. The CPE protein binds to the second extracellular loop of the claudin family of proteins (Eichner, et al. 2017) After binding to the membrane, the toxin oligomerizes resulting in the formation of hexameric pores (Robertson, et al. 2007). These cation-specific pores selectively allow the passage of Ca^{2+}. At low concentrations, CPE induces caspase-3-associated apoptosis through calpain activation, while higher concentrations of CPE result in massive Ca^{2+} influx resulting in necroptosis (Shrestha, Gohari and McClane 2019). CPE

is hypothesized to target cells in the villus tips of the small intestine causing necrosis and villus blunting (Smedley, et al. 2008).

vi. Necrotic Enteritis Beta-Like Toxin

Necrotic enteritis beta-like toxin (NetB) is a plasmid borne, pore-forming toxin associated with necrotic enteritis in avian populations, especially in chicken broiler flocks (Keyburn, et al. 2008). It is a member of the α-hemolysin family. It has 30% amino acid similarity to the alpha hemolysin produced by *S. aureus*, although it differs in its ion selectivity, pore size, and receptor (Yan, et al. 2013). The formation of unregulated ion pores resulting in death of epithelial cells appears to be the general mechanism (Rood, Keyburn and Moore 2016).

Long thought to be caused by CPA, necrotic enteritis in poultry is caused by NetB-producing strains of *C. perfringens* (toxinotype G) (Keyburn, et al. 2008). Necrotic enteritis in poultry is one of the most costly diseases caused by *C. perfringens* with an estimated economic loss of around $2 billion annually. Necrotic enteritis in poultry manifests as either an acute or a chronic form. The acute form generally occurs 2-6 weeks postpartum due to a gap in immunity caused by the depletion of maternal antibodies and the lack of a mature immune system. It results in increased mortality rates during this timeframe. The chronic form is subclinical and results in decreased growth. In both forms, lesions on the intestinal wall are observed in post-mortem necropsies (Li, Adams, et al. 2013).

1.1.2 Non-Typing Toxins

i. Perfringolysin O

Perfringolysin O (PFO) is a chromosomally encoded toxin that may be produced by any toxinotype of *C. perfringens* (Verherstraeten, Goossens, et al. 2015). It is a member of the cholesterol-dependent cytolysin family, and forms large oligomers of over 40 molecules after binding cholesterol on host cell membranes (Verherstraeten, Goossens, et al. 2015). The monomers in these large complexes insert hairpins into the cell membrane resulting in the formation of large, unregulated pores, eventually resulting in cell death (Awad, et al. 2001). While PFO itself is not essential for pathogenesis, in animal models the presence of PFO with CPA produces a synergistic effect, with PFO contributing to endothelial damage and thrombosis and the resulting tissue necrosis (Verherstraeten, Goossens, et al. 2013) (Awad, et al. 2001). PFO also has synergistic effects with ETX in mouse models (Fernandez-Miyakawa, et al. 2008).

ii. Toxin Perfringens Large

As its name suggests, Toxin Perfringens Large (TpeL) is the largest toxin produced by *C. perfringens*. It has high amino acid similarity to the *C. difficile* toxins TcdA and TcdB (39% and 38%, respectively) and to the *C. sordellii* lethal toxin (TcsL) (39%), which together form the large clostridial glycosylating toxin (LCGT) family (Amimoto, et al. 2007). The TpeL protein binds to an unknown receptor and enters into host cells by endocytosis where it changes conformation as the endosomal pH drops. This results in the catalytic domain inserting through the endosome into the cytoplasm, where it is auto-catalytically cleaved. Once free in the cytoplasm, the catalytic domain glucosylates Ras proteins using either UDP-*N*-acetyl-glucosamine or UDP-glucose,

shutting down the Ras-signaling pathway, causing apoptosis of affected cells (Nagahama, Ohkubo, et al. 2011). The presence of TpeL in *C. perfringens* Type G strains causing avian necrotic enteritis is associated with more rapid disease progression and a greater mortality rate (Coursodon, et al. 2012).

iii. Beta 2 Toxin

The function(s) of beta 2 toxin's (CPB2) are unknown; however, the presence of CPB2-encoding genes is associated with enteric diseases in animals (Bueschel, et al. 2003) (Farzan, et al. 2013). In *vitro* studies of cytotoxicity in human and porcine cells indicate that toxicity is independent of CPB2 (Allaart, et al. 2014); however, it is involved in necrosis of guinea pig intestines and CHO cells (Gibert, Jolivet-Renaud and Popoff 1997). Additional research is needed to understand how CPB2 contributes to disease progression.

iv. NetF

The non-typing NetF toxin found in some toxinotype A strains of *C. perfringens* is correlated with hemorrhagic diarrhea in canines (Leipig-Rudolph, et al. 2018) and necrotizing enteritis in neonatal foals (Gohari, et al. 2015). Canine hemorrhagic diarrhea presents with sudden onset vomiting and bloody diarrhea, and death may occur if untreated (Leipig-Rudolph, et al. 2018).

v. Additional Toxins

C. perfringens produces a large number of additional enzymes including several sialidases involved in adhesion (Li, Sayeed, et al. 2011), proteases that have been implicated in catalytic activation of other toxins (Jin, et al. 1996), and several additional chromosomally encoded toxins (Adams, et al. 2008). Binary enterotoxin (BEC), a close relative of iota toxin (43% identity), has been implicated in food born gastroenteritis

(Yonogi, et al. 2014). Continued research into the mechanisms of action of these and other newly discovered toxins in *C. perfringens* will be important to understand how these bacteria cause disease.

1.2 Non-Toxigenic Virulence Determinants

1.2.1 Agr-like Quorum Sensing

The presence of genes in *C. perfringens* having similarity to the *S. aureus* accessory-gene-regulator (Agr) system suggests a similar quorum-sensing system may regulate virulence factors in *C. perfringens. C. perfringens* contains genes encoding proteins with 50% and 46% amino acid similarity to ArgB and AgrD from *S. aureus* (Ohtani, Yuan, et al. 2009). Similar to Agr in *S. aureus*, the *C. perfringens* Agr-like system regulates both chromosomal- and plasmid-encoded toxins (Chen and McClane 2012), demonstrating its widespread regulatory function. In addition to in *vitro* regulatory functions, the Agr-like system is also required for necrotic enteritis in poultry (Yu, et al. 2017).

1.2.2 VirR/VirS System

Two component systems have important functions in signal transduction in many bacteria, including *C. perfringens.* VirR/VirS is a two-component regulatory system composed of a sensor protein (VirS), which is a histidine kinase, and a transcriptional regulator (VirR). Once activated by an external signal, the VirS sensor autophosphorylates and then transfers the phosphoryl group to the VirR regulatory protein. VirR binds to DNA, activating transcription of genes immediately downstream. VirR activity is also mediated through secondary messengers such as VirU and VirT

(Okumura, et al. 2008), as well as a VirR directed protease (Shimizu, et al. 2002). Inactivation of the VirR/VirS system attenuates the ability of the bacteria to cause disease, demonstrating its importance in pathogenesis (Ma, et al. 2011). The VirR/VirS system is also responsible for regulation of the chromosomal *pfo* and *cpa* genes (Shimizu, et al. 2002) as well as plasmid-encoded toxin genes such as *cpb2* (Ohtani, Kawsar, et al. 2003). In contrast, regulation of other toxins, such as ETX, is independent of VirR/VirS regulation (Chen, et al. 2011). VirR/VirS also regulates many non-toxin genes, with qRT-PCR analysis indicating significant regulation of 147 genes (Ohtani, Hirakawa, et al. 2010).

1.2.3 Adhesins

Historically, studies of *C. perfringens* have focused on the importance of toxins in disease progression, while minimal research has investigated the ability of the bacteria to adhere to intestinal walls. Recently, it was determined the ability of *C. perfringens* to cause necrotic enteritis in poultry correlates with the presence of putative adhesion genes (Wade, et al. 2015). In addition, type IV pili, which provide gliding motility in many clostridia (Varga, Nguyen, et al. 2006), also mediate attachment to muscle cells (Rodgers, Arvidson and Melville 2011). The presence of type IV pili is required for biofilm formation (Varga, Therit and Melville 2008). While recent research has begun to elucidate the function of adhesins in *C. perfringens* related disease, much remains to be discovered.

1.3 *C. perfringens* Genomes

C. perfringens genomes range in size from 2.69Mb to 4.17Mb (NCBI 2021). With only 12.6% of its genes considered core genes, *C. perfringens* has the most

divergent pangenome of any Gram-positive bacterial species (Kiu, Caim, et al. 2017). Contributing to this diversity is a large number of plasmids. These plasmids range in size from ~45kb to ~140 kb and may contain a number of genes coding for antibiotic resistance, toxins, or other proteins needed for pathogenesis (Li, Adams, et al. 2013). In addition, *C. perfringens* strains are able to host several plasmids at one time, which combined with the mobile elements found in plasmids creates an ideal setting for horizontal gene transfer and diversity (Li, Adams, et al. 2013).

1.4 Current Problems in *C. perfringens* Genome and Metabolic Pathway Research

Genomic research on disease causing *C. perfringens* has heavily focused on the contribution of its toxins and well established virulence factors. While some diseases are linked to the presence of particular toxins, others only seem to occur in specific circumstances, or when secondary virulence factors are present (Verherstraeten, Goossens, et al. 2015). Necrotic enteritis caused by toxinotype G is associated with the presence of the *netB* gene; however, bacteria that are unable to produce the NetB toxin also cause necrotic enteritis (Lacey, et al. 2018). In contrast, some strains that produce NetB are unable to cause the necrotic enteritis (Rood, Keyburn and Moore 2016). The inability to fulfill Koch's molecular postulates by focusing on virulence factors and their contribution to necrotic enteritis progression creates uncertainty and an inability develop effective mitigation practices (Lacey, et al. 2018). Because bacterial metabolism supplies all of the biosynthetic precursors for the synthesis of virulence factors, it is surprising there has been little research into the metabolic differences between *C. perfringens* strains. This gap in our knowledge occurs despite a large number of assembled *C. perfringens* genomes. *C. perfringens* is a strict anaerobe that is able to utilize several

carbohydrates via glycolysis and fermentation to produce acetate, ethanol, butyrate, lactic acid, carbon dioxide, and hydrogen gas (Vos, et al. 2005). In addition, *C. perfringens* is able to utilize amino acids in Stickland fermentation as another method to produce ATP (Andressen, Gottschalk and Bahl 1989). While energy production pathways in *C. perfringens* are well established, our understanding of amino acid biosynthesis is severely lacking with current genomic annotations showing as few as 45 genes involved in amino acid biosynthesis (Shimizu, et al. 2002).

1.5 Objectives of this Study

Considerable research into the contribution of toxins and other virulence factors of *C. perfringens* is available, while the contributions of metabolic genes to the disease process have largely been ignored. As a commensal organism, metabolic gene content is an important contributor of *C. perfringens*' ability to colonize animals. This thesis addresses this knowledge gap by focusing on metabolic differences in the genomes of *C. perfringens* strains that cause necrotic enteritis.

CHAPTER II MATERIALS AND METHODS

2.1 Strain Selection

C. perfringens genomes were selected based on availability of metadata information and the quality of genomic data present on the Genome database at the National Center for Biotechnology Information (NCBI). The inclusion criteria were: 1) to have complete genome assembly; 2) to contain both *cpa* and *netB*; and 3) the NCBI/GenBank metadata for each genome was examined for the ability to cause necrotic enteritis. The completeness of genome assembly is assessed using several metrics in the Pathosystems Resource Integration Center (PATRIC) (Wattam, et al. 2014). Specifically, CheckM (Parks, et al. 2015) completeness scores of 100 and contamination scores <3 (indicating very low contamination) were chosen. For the chosen genomes, the PATRIC EvalG and EvalCon annotation analysis tools (Parrello, et al. 2019) showed coarse consistency scores ≥99.7 (indicating presence of proteins predicted by EvalG) and fine consistency scores ≥98 (indicating correct numbers of protein roles predicted by EvalCon). The very high coarse consistency as well as fine consistency scores indicate the genomes to be nearly complete and contain very little (or no) contamination, confirming the genome selection criteria used. The genomic sequence of *C. perfringens* strain ATCC 13124 was also included as a reference. In total, 19 genomes were selected for this study. The accession numbers and other information of the genomic sequences used in this study are listed in Table 3, with additional information in Appendix A.

Table 3: Genomic sequences used in this study and associated information

STRAIN NAME	ACCESSION NUMBER	COUNTRY OF ISOLATION	YEAR OF ISOLATION	GENOME SIZE (MB)	PATRIC ANNOTATION CDS COUNT
ATCC13124	CP000246.1	N/A		3.26	2866
STRAIN 37	PJSN01000000.1	Belgium	2005	3.48	3229
STRAIN 67	PJSM01000000.1	Belgium	2005	3.45	3211
STRAIN 98.787 18-2	PJTE01000000.1	Denmark	2005	3.61	3433
DEL1	CP019576.1	USA	2009	3.81	3679
EHE-NE18	CP025501.1	Australia	2002	3.66	3398
EHE-NE7	PJTC01000000.1	Australia	2002	3.43	3199
EUR-NE15	PJTB01000000.1	Australia	2002	3.58	3377
FC2	PJTA01000000.1	USA	1995	3.73	3272
GNP-1	PJSZ01000000.1	USA	1996	3.73	3621
ITX1105-12MP	PJSY01000000.1	USA	1996	3.65	3518
PENNINGTON	PJSQ01000000.1	USA	1993	3.75	3655
SOM-NE35	PJTM01000000.1	Australia	2007	3.57	3339
TAM-NE38	PJTL01000000.1	Australia	2011	3.65	3527
TAM-NE40	PJSK01000000.1	Australia	2012	3.65	3561
TAM-NE46	PJSI01000000.1	Australia	2013	3.68	3655
UDE 95-1372	PJTI01000000.1	USA	1995	3.64	3515
WARREN	PJTH01000000.1	Australia	1993	3.41	3648
WER-NE36	PJTG01000000.1	USA	2010	3.78	3279
***C. VENTRICULI* STRAIN 17**	BCMW00000000.1	Japan	2014	2.47	2325

2.2 Phylogenetic Analysis

A phylogenetic tree was constructed using *Clostridium ventriculi (née Sarcina ventriculi*) strain 17 as an outgroup due to its close relation to *C. perfringens* (Cruz-Morales, et al. 2019). Table 3 includes the accession number and other information for *C. ventriculi* strain 17. The phylogenetic analysis was performed using the 'Codon Tree' method available in PATRIC (https://patricbrc.org/app/PhylogeneticTree). PATRIC identifies protein-coding genes from each genome based on their own protein family (PGFams) database (Davis, et al. 2016). A set of single-copy genes encoding PGFams found across the given set of genomes are used to generate protein sequence alignments using Multiple Alignment using fast Fourier Transform (MAFFT) (Katoh, et al. 2002). Coding nucleotide sequence alignments are generated based on the protein sequence

alignments. These alignments are concatenated and used with the Randomized Axelerated Maximum Likelihood (RAxML) version 8.2.11 (Stamatakis 2014) to reconstruct a maximum likelihood phylogenetic tree with 100 bootstrap replications. The *C. ventriculi* strain 17 genome was used as an outgroup. The PATRIC pipeline identified 739 single-copy genes shared across the 20 genomes (19 *C. perfringens* genomes and the outgroup). Strain genomes were put into clades based on their phylogenetic relationship. The details for the phylogenetic reconstruction (command lines used and statistics) are found in Appendix B. The reconstructed phylogeny including the outgroup is found in Figure A1.

2.3 Visualization of Ortholog Distribution among the Genomes

Circos (Krzywinski, et al. 2009) was used to create a circular visualization of the orthologous gene distribution among *C. perfringens* genomes. *C. perfringens* ATCC13124 genome was used as the reference. One genome from each of the six clades (see 3.2) was selected. A Reciprocal Best Hits (RBH) search using blastp (Altschul, et al. 1990) was performed between the reference and each representative genomes to identify CDS's in the selected genomes orthologous to those in the reference genome.

2.4 Comparative Proteomic Analysis

Proteome comparisons were performed using genus-specific protein families (PLFams) identified using PATRIC (Davis, et al. 2016). The presence/absence of protein families across genomes were examined by comparing PLFams found from each genome. The PLFams that were uniquely present in each genome as well as those that were uniquely absent were identified.

2.5 Metabolic Pathway Analysis

The PathoLogic tool in Pathway Tools V25.0 (Karp, et al. 2021) was utilized to generate pathway databases for each genome based on the NCBI annotation. The pathway databases were generated using taxonomic pruning enabled to prevent false positives and a default pathway prediction cutoff score of 0.15. Pathway Tools V25.0 was also used to assess which pathways were present or absent uniquely or commonly among the genomes. In each case where a predicated pathway was absent, a tblastn similarity search (Altschul, et al. 1990) using the protein sequence from a genome that was correctly identified as the query against the translated genomic sequences was performed to search the genomes where the pathway was not identified. This identified several cases where annotations were incorrect or a pseudogene was identified.

CHAPTER III RESULTS AND DISCUSSION

3.1 Selection of *C. perfringens* Strains and Genomes

The 19 *C. perfringens* genomes selected for analysis are listed in Table 2. The genome size ranges from 3.26Mb to 3.81Mb with an average length of 3.59Mb. The average number of protein-coding sequences (CDS) in the PATRIC annotation was 3,420 with a range of 2,866 to 3,679 while the average number of CDS in the NCBI annotation was lower at 3,269 with a range of 2,876 to 3,489. Additional information about the selected strains is found in Appendix A

3.2 Phylogenetic Relationships of Selected *C. perfringens* genomes

The phylogenetic tree of the 19 *C. perfringens* genomes was first reconstructed using *C. ventriculi* as an outgroup (Figure B1). Figure 1 shows the same phylogeny without including the outgroup. This phylogeny showed that the 19 *C. perfringens*

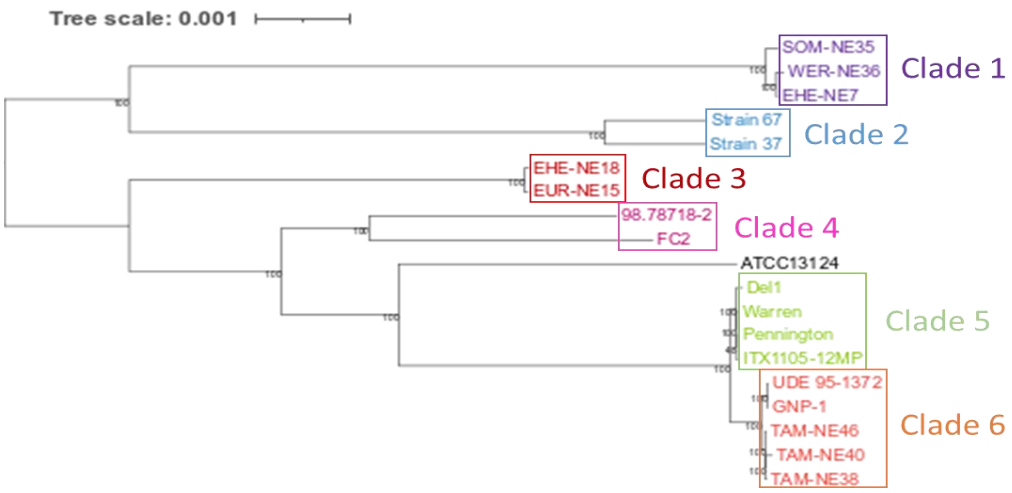

Figure 1: Phylogenetic reconstruction of selected *C. perfringens* strain genomes.

The original phylogeny is shown in Figure B1. The outgroup (*C. ventriculi*) was removed for the visual clarity. The genomes are grouped into six clades (Clade 1: purple, Clade 2: blue, Clade 3: red, Clade 4: magenta, Clade 5: green, and Clade 6: orange).

genomes are divided into six clades. Only one clade (Clade 2) consists of genomes from the same country of origin, where strain 37 and strain 67 were both isolated in Belgium during the same study (Lacey, et al. 2018). The strain genomes from the United States, Australia, and Denmark revealed little grouping based on country of origin.

3.3 Distribution of Orthologs in *C. perfringens* genomes

The circular visualization of ortholog distribution among the *C. perfringens* genomes is presented in Figure 2. Several orthologous CDS's with high variability are marked in red. Region A encompassing nucleotide positions 128,962 through 169,326 of the reference genome codes for several ABC transporter related proteins and proteins with homology to *E. coli* telluride resistance (*ter*). The second region (B; 473,690 through 488,405) contains fourteen CDS coding proteins involved in the uptake and utilization of hyaluronate. Region C, which is the largest, stretching from nucleotide positions 1,073,505 through 1,234,200 of the reference genome contains many hypothetical proteins, although it does contain many phage related genes and genes related to fucose utilization. Similar to region C, region D (1,783,902 through 1,820,416) contains a large number of hypothetical proteins with many phage related genes interspersed throughout. Finally, region E (2,857,812 through 2,876,059) contains many genes associated with Type 1 recombinase systems. Regions A and B contain unique telluride resistance and metabolic genes not found in the genomes selected from each clade, while regions C, D, and E show many CDS's indicative of phage insertions that are not present in the selected strains.

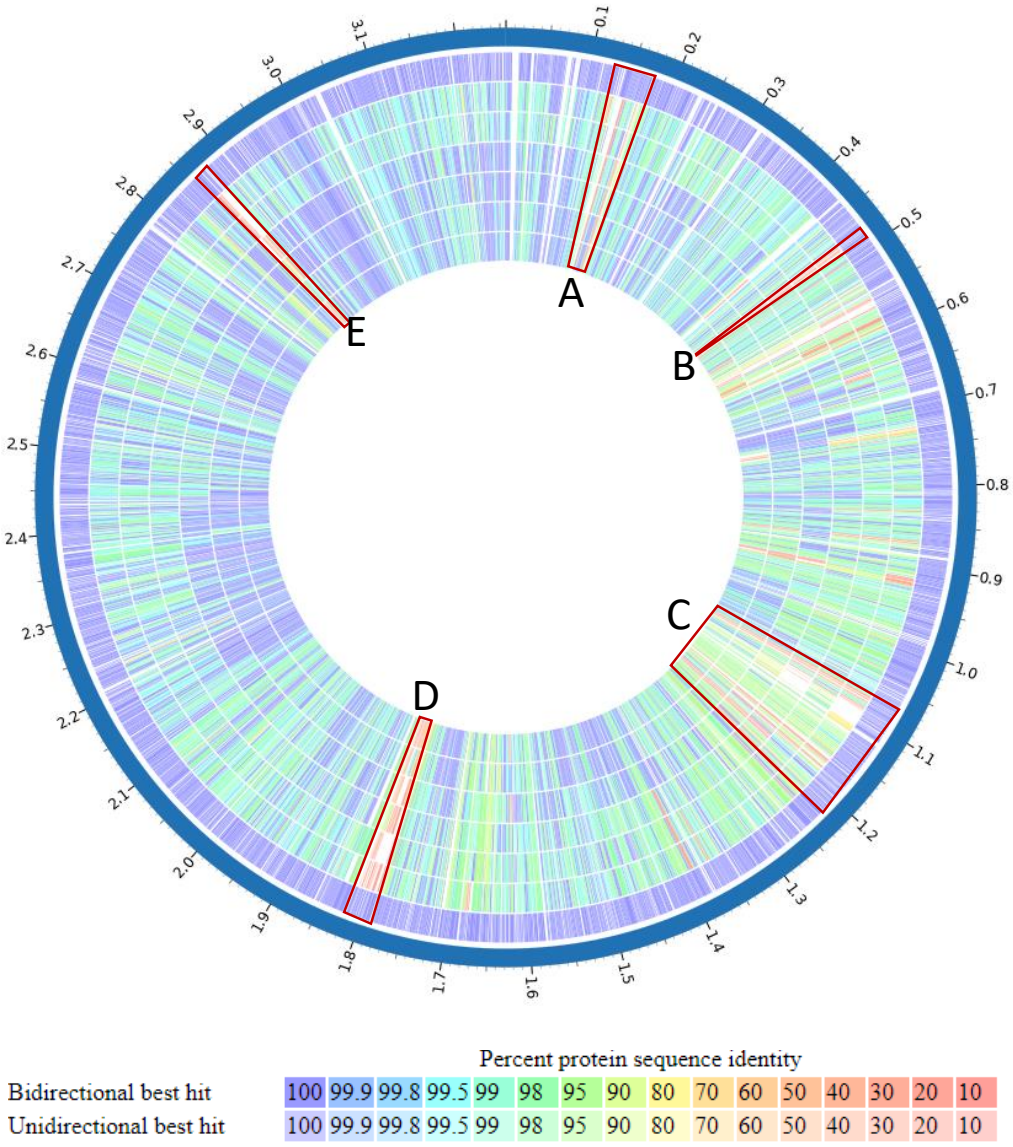

Figure 2: Ortholog distribution among selected *C. perfringens* genomes.

The reference genome (ATCC 13124) is on the outermost track. Orthologous CDS's identified from six genomes representing the six clades are on the inner tracks with color coded % sequence identity (from outside to inside: Strain 37, EHE-NE7, EUR-NE15, 98.78718-2, Del1, GNP-1). Regions with orthologous CDS's with high sequence divergence from the reference are marked with red boxes (A-E). The genomic positions (in Mb) shown on the outer circle are based on the reference genome. Note that the CDS locations are shown only for the reference genome. For the other six genomes, colored boxes indicate only presence and absence of CDS's orthologous to those on the reference genome.

3.4 Comparison of Protein Families among *C. perfringens* Genomes

A total of 4,415 PLFams protein families were identified across the 19 *C. perfringens* genomes with 2,454 present in all genomes. While the majority of these uniquely present or absent families are annotated as "hypothetical" without any function associated, several of the better characterized protein families have unique distribution patterns. Strain FC2 uniquely contains proteins associated with a defined CRISPR family, which are lacking in the other 18 genomes. Long, et al. (2019) identified CRISPR systems in all but four of the genomes selected for this study. However, the FC2 genome was the only one with proteins associated with a CRISPR system belonging to a defined family (Long, et al. 2019). The genomes of strains 37, 67, 98.78718-2, and EHE-NE18 contained several unique phage related family proteins. Three genomes (*i.e.*, Pennington, Warren, and UDE 95-1372) contained lincosamide nucleotidyltransferases, although they were of two different types; the Pennington and Warren genomes contained a Lnu(D) type, while the UDE 95-1372 genome contained a Lnu(P) type. The TAM-NE46 genome had several unique protein families associated with phosphotransferase (PTS) systems. It contained unique proteins in the mannose-specific IIB component, ascorbate-specific IIA component, and cellobiose-specific IIC protein families. The ATCC 13124 genome uniquely carried proteins placed in the hyaluronate-oligosaccharide-specific IIA, IIB, IIC and IID families, which were not present in any other genomes.

3.5 Metabolic Pathway Analysis among *C. perfringens* Genomes

Using Pathway Tools, metabolic pathways were identified in the 19 *C. perfringens* genomes. Comparison of the metabolic pathway repertoire among these genomes revealed interesting biochemical differentiation among *C. perfringens* strains. Figure 3 summarizes the pathways where the unique distribution among the 19 genomes was identified.

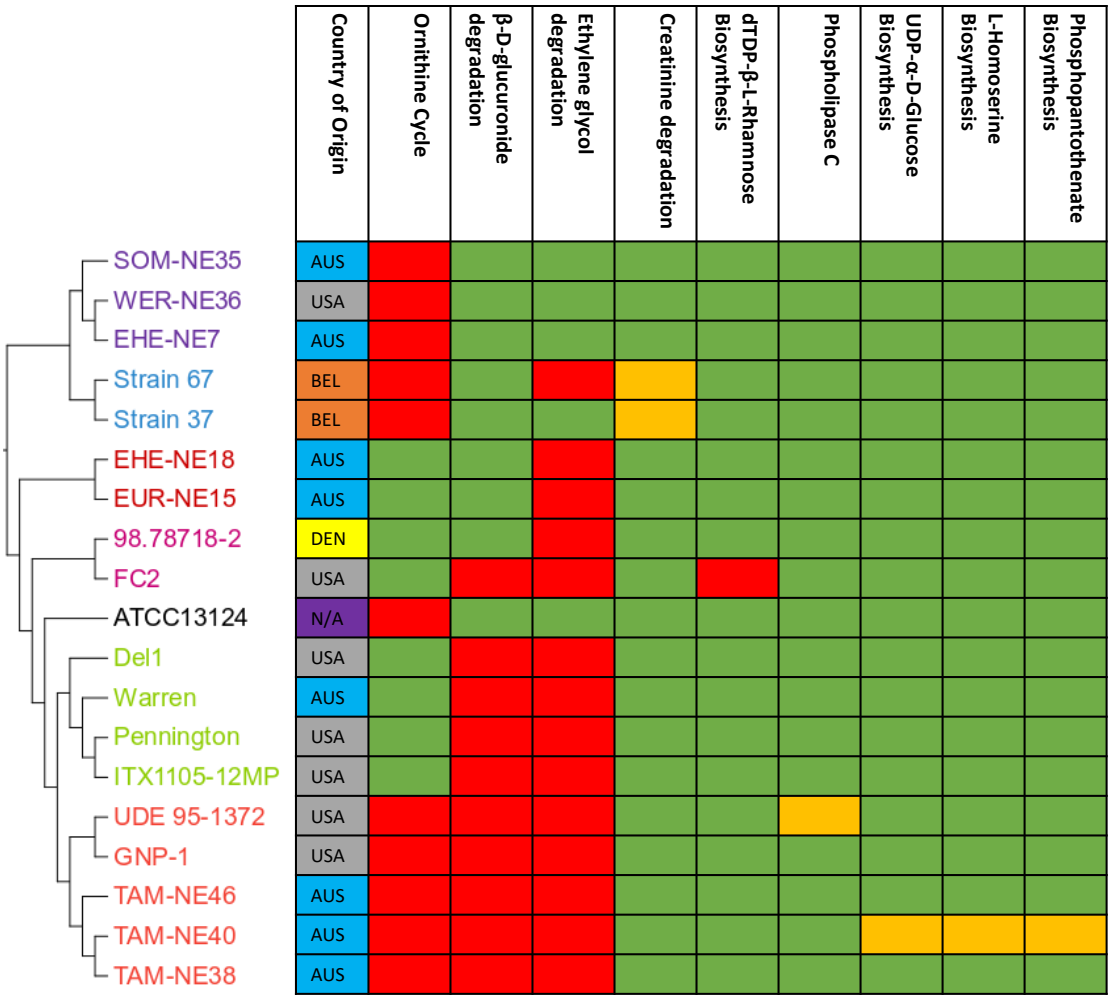

Figure 3: Metabolic pathway distribution among the 19 *C. perfringens* genomes.

Phylogenetic relationship on the left is derived from Figure 1. The table shows the presence (green) or absence (red) of a functional pathway. When pseudogenes are found for the enzymes involved in the pathway, it is indicated with orange.

3.5.1 Carbohydrate Biosynthesis Pathways

Two of the carbohydrate biosynthetic pathways (UDP-α-D-glucose biosynthesis and dTDP-β-L-rhamnose biosynthesis) were not recognized in two genomes, each missing a different pathway. The TAM-NE40 genome appears to lack the UDP-α-D-glucose biosynthesis pathway shown in Figure 4 because a phospho-sugar mutase found in all other genomes was not identified. In the NCBI annotation of the TAM-NE40 genome, the gene sequence of the phospho-sugar mutase (CYK79_08245) in fact exists in the correct location with the same gene order as other genomes. However, the gene region included a frame-shifting deletion, resulting in a truncation of the CDS (Figure 5), and was annotated as a pseudogene. The truncated protein could lead to a metabolic block in UDP-α-D-glucuronate biosynthesis, which would hinder capsule synthesis (Kalelkar, et al. 1997). The FC2 genome is the only one that lacks the dTDP-β-L-rhamnose biosynthesis pathway. As shown in Figure 6, it involves four enzymes: glucose-1-phosphate thymidylyltransferase, dTDP-glucose 4,6-dehydratase, dTDP-4-dehydrorhamnose 3,5-epimerase, and dTDP-4-dehydrorhamnose reductase. All other genomes contain all genes encoding these four enzymes in a single cluster. The genomic region containing these genes may have been lost. Alternatively sequencing/assembly failure may have caused the lack of these genes in strain FC2.

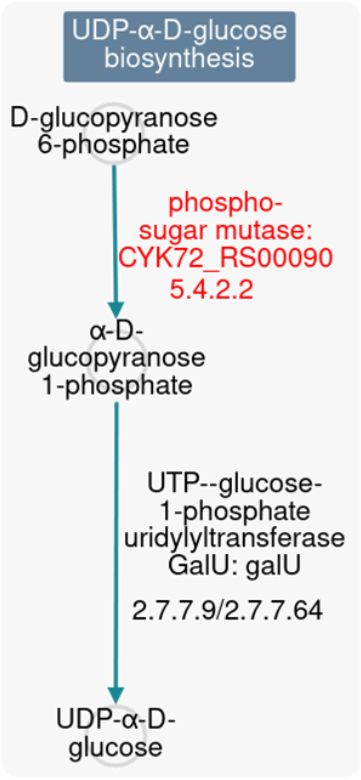

Figure 4: UDP-α-D-glucose biosynthesis pathway.

The metabolites and enzymes involved in the pathway are shown. The enzyme missing from the TAM-NE40 genome is shown in red. The pathway figure was obtained from Pathway Collages (Paley, et al. 2016).

```
          1414                                                                1476
          K  C  I  D  G  L  R  N  D  A  L   K  E  M  N  G  V  K  V  I  T
TAM-NE38  aagtgcatagatggcttaagaaatgatgccc-taaaagaaatgaatggagttaaagttattaca
TAM-NE40  aagtgcatagatggcttaagaaatgatgcccctaaaagaaatgaatggagttaaagttattaca
          K  C  I  D  G  L  R  N  D  A  P  K  R  N  E  W  S  *  S  Y  Y  I
           S  A  *  M  A  *  E  M  M  P  L  K  E  M  N  G  V  K  V  I  T
            V  H  R  W  L  K  K  *  C  P  *  K  K  *  M  E  L  K  L  L  H
```

Figure 5: Frame-shifting insertion found in the phospho-sugar mutase gene in TAM-NE40.

Part of the sequences of the phopho-sugar mutase gene is shown. The functional sequence from TAM-NE38 (top) is aligned with the pseudogene sequence found in TAM-NE40, which includes an insertion of an extra cytidine in a poly-c region resulting in an early stop codon. The blue indicates the normal amino acid coding seen in TAM-NE38, while the red lettering shows where the CDS differs and terminates at the premature stop.

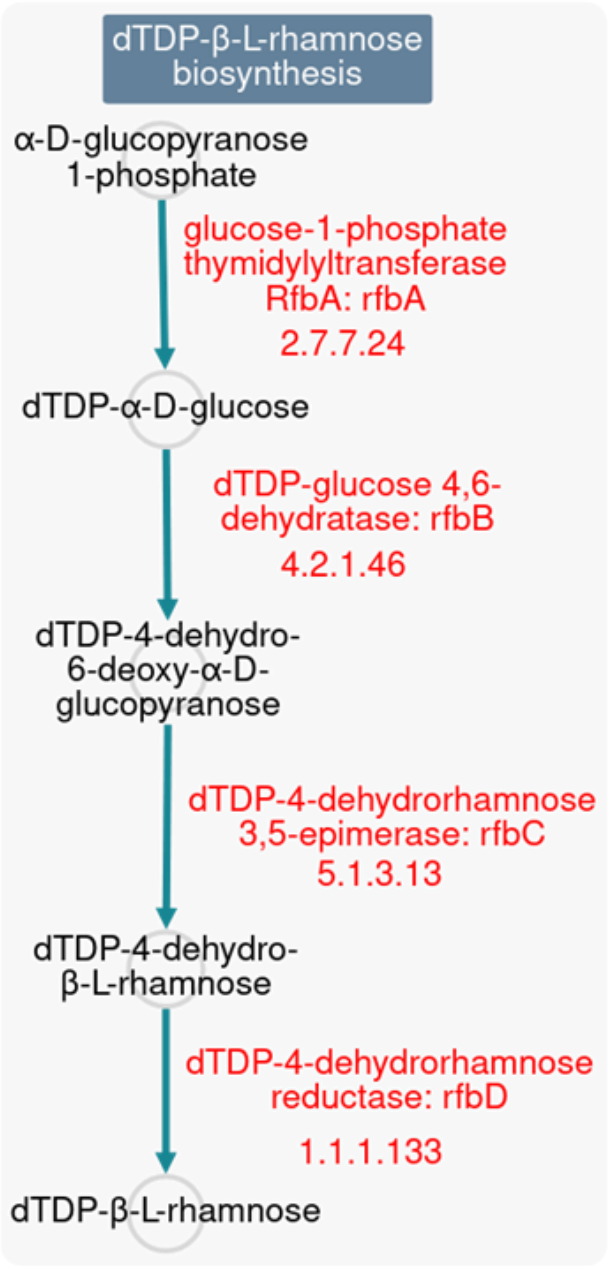

Figure 6: dTDP-β-L-rhamnose biosynthesis pathway.

The metabolites and enzymes involved in the pathway are shown. The enzymes missing from the FC2 genome are shown in red. The pathway figure was obtained from Pathway Collages (Paley, et al. 2016).

3.5.2 Ornithine Cycle

As shown in Figure 3, the complete ornithine cycle is lacking in eleven of the nineteen genomes. This is due to the lack of the gene encoding arginase in these genomes (Figure 7). Absence of the arginase gene in these genomes was confirmed by a tblastn similarity search of the genomes in which it was not annotated. While clades 1, 2, 6 and the reference strain lack the arginase, clades 3, 4, and 5 contain the arginase genes.

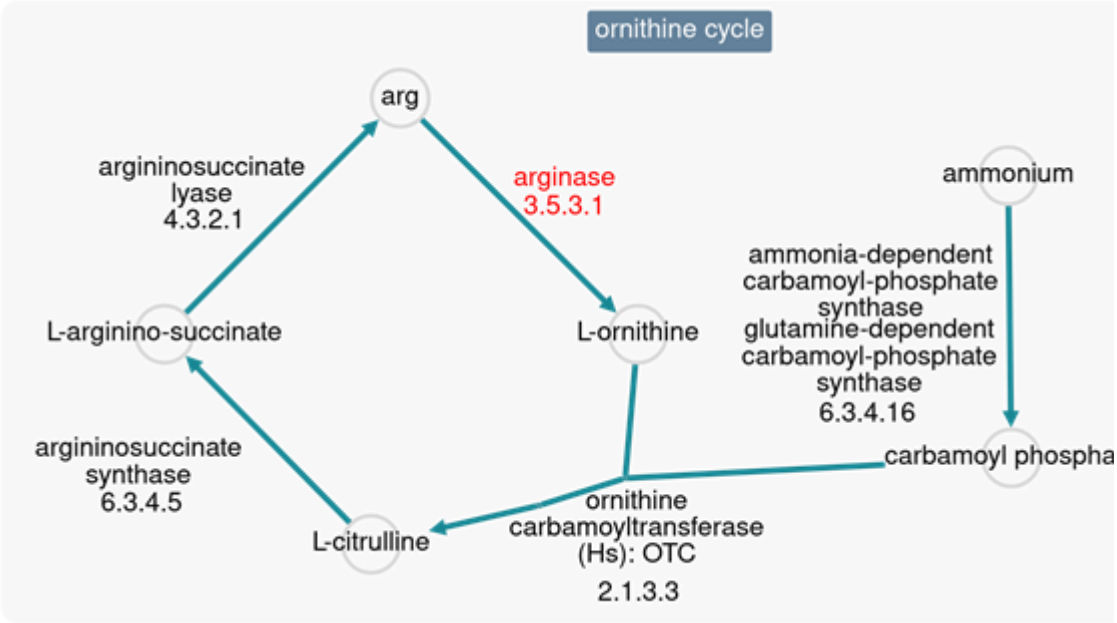

Figure 7: Ornithine cycle.

The metabolites and enzymes involved in the ornithine cycle are shown. The arginase in red is missing from ten of the selected genomes. The pathway figure was obtained from Pathway Collages (Paley, et al. 2016).

3.5.3 Phospholipase C Pathway

As described in 1.2.1, alpha toxin (CPA) possesses phospholipase activity (Sakurai, Nagahama and Oda 2004); hence, the phospholipase C pathway should be

found in *C. perfringens*. Despite this expectation, the UDE 95-1372 genome appears to lack a functional phospholipase C pathway. The coding region of the *cpa* gene (CYK86_14610) in the UDE 95-1372 genome has a nonsense mutation changing the 124th codon from a serine (tca) to a stop (taa) (Figure 8), and is annotated as a pseudogene. As all known strains of *C. perfringens* produce the CPA protein, an inactivated *cpa* gene would be an unusual evolutionary event.

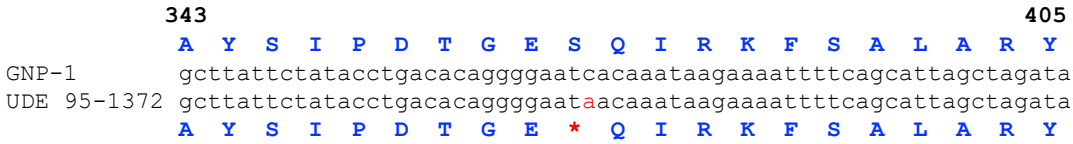

Figure 8: Nonsense mutation found in the *cpa* gene in UDE-95-1372.

Part of the sequences for *cpa* gene in UDE 95-1372 (bottom) and its closest relative GNP-1 (top) are aligned. The point mutation at position 372 (from c to a, shown in red) in UDE-95-1372 causes the 124th amino acid to change from serine to a stop codon.

3.5.4 Carboxylate Degradation Pathways

As shown in Figure 3, the β-D-glucuronide degradation pathway (part of the carboxylate degradation class) separates the 19 *C. perfringens* genomes into two groups: those that have the β-D-glucuronide and D-glucuronate degradation (Figure 9) and D-fructuronate degradation (Figure 10) pathways and those without. The genomes lacking these pathways are all missing the same four genes. Two genes missing from the β-D-glucuronide and D-glucuronate degradation pathway are those encoding β-glucuronidase and glucoronate isomerase. From the D-fructuronate catabolism pathway both mannonate dehydratase and bifunctional 4-hydroxy-2-oxoglutarate aldolase/2-dehydro-3-deoxy-phosphogluconate aldolase are missing. In each case, a tblastn search resulted in no

matches. Clades 5 and 6 were universally missing all four of these genes while FC2 in clade 4 was also missing them. Clades 5 and 6 are phylogenetically related. Therefore, these enzymes must have been lost before the divergence of these clades. On the contrary, lineage-specific loss appears to have happened in FC2 since another strain belonging to clade 4 contains all of these enzymes.

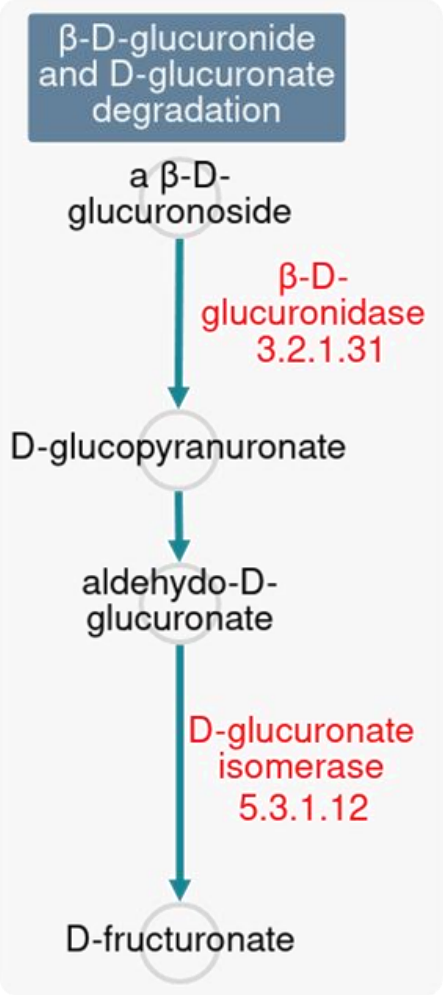

Figure 9: β-D-glucuronide and D-glucuronate degradation pathway.

The metabolites and enzymes involved in the pathway are shown. The enzymes in red are missing from the ten genomes. The pathway figure was obtained from Pathway Collages (Paley, et al. 2016).

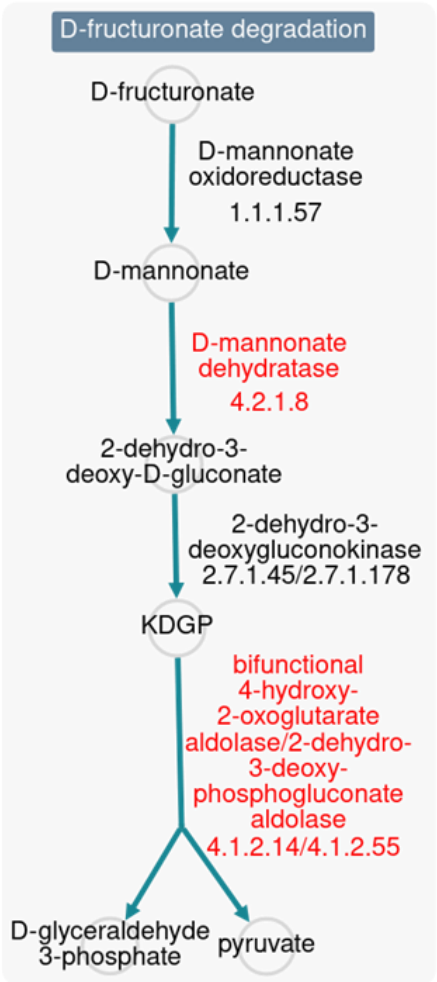

Figure 10: D-fructuronate degradation pathway.

The metabolites and enzymes involved in the pathway are shown. The enzymes in red are missing from the ten genomes. The pathway figure was obtained from Pathway Collages (Paley, et al. 2016).

3.5.5 Amino Acid Biosynthesis Pathways

The amino acid synthesis pathways appear to be relatively stable across these *C. perfringens* strains. Of the 19 genomes, one genome (i.e., TAM-NE40) lacked support for the L-homoserine biosynthesis pathway. The coding region of the aspartate kinase gene (CYK79_07045) in the TAM-NE40 genome has a frame-shifting deletion in the 65th codon causing premature termination, and is annotated as a pseudogene (Figure 11).

Figure 11: Alignment of the aspartate kinase gene in TAM-NE40.

Shows the sequence for the aspartate kinase identified as part of the L-homoserine biosynthesis pathway. TAM-NE46 (Top) has a functional aspartate kinase, while TAM-NE40 shows a deletion of an adenine in a poly-a region resulting in a frameshift into an early stop codon.

3.5.6 Amine and Polyamine Degradation Pathways

Three pathways in the Amine and Polyamine Degradation class were present in some *C. perfringens* genomes. Two pathways were identified in all genomes (N-acetylglucosamine degradation I and ethanolamine utilization). While the creatinine degradation pathway (Figure 12) was present in most genomes (Figure 3) the genomes of strains 37 and 67 appeared to lack the creatininase gene. In both of these genomes, the creatininase coding regions are identified and annotated correctly (CYK73_06140 and CYK75_06295, respectively). However, both of these coding regions have the same single C-to-A nucleotide substitution at (Figure 13), resulting in a premature stop codon

truncating the predicted protein by 27 amino acids. For this reason, these genes were annotated as pseudogenes. This is a surprising observation because creatininase hydrolyzes creatinine to creatine that can be phosphorylated by creatine kinase to creatine-phosphate, a small phospho donor used in substrate-level phosphorylation. As shown in Figure 13, in both strains, the nucleotide sequences after the premature stop codons are conserved (up to the original stop codons). Therefore, it is likely that this substitution happened very recently just before the divergence of these two genomes. It will be interesting to determine the effect of the truncation on protein function.

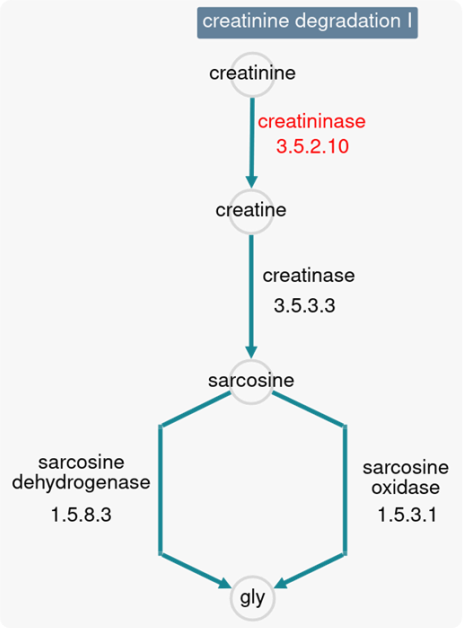

Figure 12: Creatinine degradation pathway.

The enzymes and metabolites involved in the pathway are shown. The creatininase shown in red is non-functional in strain 37 and strain 67. The pathway figure was obtained from Pathway Collages (Paley, et al. 2016).

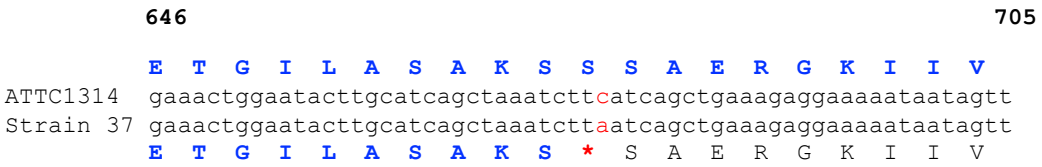

Figure 13: Truncation of creatininase in Strain 37 and Strain 67.

Part of the sequences for the creatininase are shown. ATTC 13142 (top) has a functional creatininase, while strain 37 and strain 67 (bottom) show a cytidine to adenine substitution resulting in an early stop codon (red).

3.5.7 Alcohol Degradation Pathways

The majority of pathways in the Alcohol Degradation pathway class had consistent support in all 19 genomes examined. The exception was the ethylene glycol degradation pathway (Figure 14). As shown in Figure 3, the lactaldehyde reductase gene was found in five genomes, hence the ethylene glycol degradation pathway is considered missing in all other genomes. Protein similarity search using the lactaldehyde reductase protein sequence from ATCC 13124 as the query was performed to confirm the existence or absence of the lactaldehyde reductase gene in each genome. The lactaldehyde reductase protein sequences are highly conserved and the sequences are almost 100% identical to each other. For those that lack the lactaldehyde reductase gene, the most similar sequences were other members of the iron-containing alcohol dehydrogenase (Fe-ADH) family such as L-threonine dehydrogenase and 1,3-propanediol dehydrogenase, with amino acid similarities ranging from 60% to 40%. This confirmed the presence of the lactaldehyde reductase gene only in five of the nineteen genomes. The genomes of clade 1 all had lactaldehyde reductase, while clade 2 had strain 67 lacking the gene and strain 37 containing it.

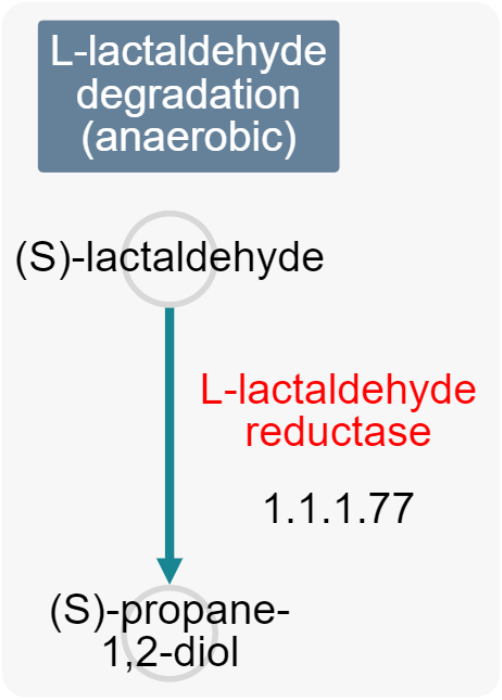

Figure 14: L-lactaldehyde degradation pathway.

The enzymes and metabolites involved in the pathway are shown. L-lactaldehyde reductase (red) is only present in five genomes. The pathway figure was obtained from Pathway Collages (Paley, et al. 2016).

3.5.8 Phosphopantothenate Biosynthesis 1 Pathways

Among the Biosynthesis – Cofactor, Carrier, and Vitamin Biosynthesis pathway class only one pathway had a difference. The phosphopantothenate biosynthesis I pathway was not identified in the TAM-NE40 genome (Figure 3). This is because the gene encoding the 2-dehydropantoate 2-reductase enzyme was not found in this genome. Inspection of the genome showed the gene exists (CYK79_01055) but was annotated as a pseudogene. As shown in Figure 15, a single nucleotide insertion at the position 283 resulted in a frame-shift causing changes in the amino acid sequence and a stop codon at the positions 322-324.

Figure 15: Frame-shifting mutation and truncation of 2-dehydropantoate 2-reductase in TAM-NE40.

Part of the sequences for the 2-dehydropantoate 2-reductase gene is shown. TAM-NE46 (top) has a functional 2-dehydropantoate 2-reductase (blue amino acid sequence), while TAM-NE40 (bottom) shows an insertion in a poly-A region causing a frame-shift (red) resulting in a premature stop codon.

3.6 Discussion

Comparative genomic analysis among the 19 *C. perfringens* genomes revealed differences in metabolic pathways between some genomes. Most pathways examined were either present or absent by the clade, while there were instances of strains showing genotypes not in line with their clade (most notably FC2 and TAM-NE40). Importantly, the phylogeny was reconstructed using the single-copy genes that are shared across all 19 genomes and was not affected by the presence or absence of particular genes. Therefore, phylogenetic grouping of six clades was not affected by the differential presence and absence of the genes. In addition, the presence of metabolic pathways appears to be independent of geography, the exception being the strains from Belgium (*i.e.*, strain 37 and strain 67) having a common mutation in their creatininase genes. Some genomic

differences identified in this study are undoubtedly due to errors in sequencing and/or annotation. For instance, the genome of strain TAM-NE40 had three pseudogenes that appear to disrupt pathways. This was not the case in any other strains. All of the pseudogenes in TAM-NE40 were caused by insertions or deletions in poly-C or poly-A regions. Therefore, it is possible that they are the results of sequencing or assembly errors with the genome rather than actual mutations. There were also instances when a pathway was marked as not present by PathoLogic, but the missing genes were found using a tblastn similarity search of the genome. This most often occurred due to annotations not identifying a particular gene. This underscores the importance of high-quality genome annotation to utilize the growing sequencing data for functional analysis.

CHAPTER IV CONCLUSIONS AND FUTURE STUDIES

C. perfringens is an extremely versatile pathogen that is able to cause a wide range of diseases. While virulence factors are common determinants of disease processes, metabolism is an often overlooked contributor to pathogenesis. Metabolic differences among strains could play a key role in ability of a strain to colonize the gut of infected organisms and cause a disease. The toxinotype G strains of *C. perfringens* used in this study had marked metabolic differences despite causing the same disease in host animals. The use of genomic data to identify the metabolic abilities of strains facilitates better understanding of these pathogens; however, it also comes with drawbacks. As the quantity of available genomes grows, the curation of data is important. Properly curated data will allow the modern tools being developed, many of which are based on machine learning, to effectively and more accurately process the data being produced. Verification of genome annotations will be a valuable next step in the field. Evaluating the impact of these metabolic differences on the pathogen's ability to cause diseases and the severity of diseases could also help increase understanding what factors are necessary to cause clinical diseases instead of the common sub-clinical infection.

REFERENCES

Adams, Jarrett J., Katie Gregg, Edward A. Bayer, Alisdair B. Boraston, and Steven P. Smith. 2008. "Structural basis of Clostridium perfringens toxin complex formation." *Proceedings of the National Academy of Sciences of the United States of America* 105 (34): 12194-12199. doi:10.1073/PNAS.0803154105.

Adams, Vicki, Xiaoyan Han, Dena Lyras, and Julian I. Rood. 2018. "Antibiotic resistance plasmids and mobile genetic elements of Clostridium perfringens." *Plasmid* 99: 32-39. doi:10.1016/J.PLASMID.2018.07.002.

Allaart, Janneke G., Alphons J.A.M. van Asten, Johannes C.M. Vernooij, and Andrea Gröne. 2014. "Beta2 toxin is not involved in in vitro cell cytotoxicity caused by human and porcine cpb2-harbouring Clostridium perfringens." *Veterinary Microbiology* 171 (1): 132-138. doi:10.1016/J.VETMIC.2014.03.020.

Altschul, S F, W Gish, W Miller, E W Myers, and D J Lipman. 1990. "Basic local alignment search tool." *Journal of Molecular Biology* 215 (3): 403-410. doi:10.1016/S0022-2836(05)80360-2.

Amimoto, Katsuhiko, Taichi Noro, Eiji Oishi, and Mitsugu Shimizu. 2007. "A novel toxin homologous to large clostridial cytotoxins found in culture supernatant of Clostridium perfringens type C." *Microbiology* 153 (4): 1198-1206. doi:10.1099/MIC.0.2006/002287-0.

Andressen, J R, G Gottschalk, and H Bahl. 1989. *Introduction to the Physiology and Biochemistry of the Genus Clostridium In: Minton N.P., Clarke D.J. (eds) Clostridia. Biotechnology Handbooks.* Vol. 3. Boston, MA: Springer.

Awad, Milena M., Darren M. Ellemor, Richard L. Boyd, John J. Emmins, and Julian I. Rood. 2001. "Synergistic Effects of Alpha-Toxin and Perfringolysin O in Clostridium perfringens-Mediated Gas Gangrene." *Infection and Immunity* 69 (12): 7904-7910. doi:10.1128/IAI.69.12.7904-7910.2001.

Aziz, Ramy K, Daniela Bartels, Aaron A Best, Mather DeJongh, Terrence Disz, Roberts A Edwards, Kevin Formsma, et al. 2008. "The RAST Server: Rapid Annotations using Subsystems Technology." *BMC Genomics* 9 (1). doi:10.1186/1471-2164-9-75.

Billington, Stephen J., Eva U. Wieckowski, Mahfuzur R. Sarker, Dawn Bueschel, J. Glenn Songer, and Bruce A. McClane. 1998. "Clostridium perfringens Type E Animal Enteritis Isolates with Highly Conserved, Silent Enterotoxin Gene Sequences." *Infection and Immunity* 66 (9): 4531-4536. doi:10.1128/IAI.66.9.4531-4536.1998.

Briggs, David C., Claire E. Naylor, James G. Smedley, Natalya Lukoyanova, Susan Robertson, David S. Moss, Bruce A. McClane, and Ajit K. Basak. 2011. "Structure of the food-poisoning Clostridium perfringens enterotoxin reveals similarity to the aerolysin-like pore-forming toxins." *Journal of Molecular Biology* 413 (1): 138-149. doi:10.1016/J.JMB.2011.07.066.

Bueschel, Dawn M., B. Helen Jost, Stephen J. Billington, Hien T. Trinh, and J. Glenn Songer. 2003. "Prevalence of cpb2, encoding beta2 toxin, in Clostridium perfringens field isolates:

correlation of genotype with phenotype." *Veterinary Microbiology* 94 (2): 121-129. doi:10.1016/S0378-1135(03)00081-6.

Chen, Jianming, and Bruce A. McClane. 2012. "Role of the Agr-Like Quorum-Sensing System in Regulating Toxin Production by Clostridium perfringens Type B Strains CN1793 and CN1795." *Infection and Immunity* 80 (9): 3008-3017. doi:10.1128/IAI.00438-12.

Chen, Jianming, Julian I. Rood, and Bruce A. McClane. 2011. "Epsilon-Toxin Production by Clostridium perfringens Type D Strain CN3718 Is Dependent upon the agr Operon but Not the VirS/VirR Two-Component Regulatory System." *Mbio* 2 (6): 1-11. doi:10.1128/MBIO.00275-11.

Coursodon, C.F., R.D. Glock, K.L. Moore, K.K. Cooper, and J.G. Songer. 2012. "TpeL-producing strains of Clostridium perfringens type A are highly virulent for broiler chicks." *Anaerobe* 18 (1): 117-121. doi:10.1016/J.ANAEROBE.2011.10.001.

Cruz-Morales, Pablo, Camila A Orellana, George Moutafis, Glenn Moonen, Gonzalo Rincon, Lars K Nielsen, and Esteban Marcellin. 2019. "Revisiting the Evolution and Taxonomy of Clostridia, a Phylogenomic Update." *Genome Biology and Evolution* 11 (7): 2035-2044. doi:https://doi.org/10.1093/gbe/evz096.

Davis, James, Svetlana Gerdes, Gary Olsen, Robert Olson, Gordon Pusch, Maulik Shukla, Veronika Vonstein, Alice R Wattam, and Hyunseung Yoo. 2016. "PATtyFams: Protein families for the microbial genomes in the PATRIC database." *Frontiers in Microbiology.* doi:10.3389/fmicb.2016.00118.

Diancourt, Laure, Jean Sautereau, Alexis Criscuolo, and Michel Popoff. 2019. "Two Clostridium perfringens Type E Isolates in France." *Toxins* 11 (3): 138. doi:10.3390/TOXINS11030138.

Dorca-Arevalo, Jonatan, Eduard Dorca, Benjamin Torrejon-Escribano, Marta Blanch, Mireia Martin-Satue, and Juan Blasi. 2020. "Lung endothelial cells are sensitive to epsilon toxin from Clostridium perfringens." *Veterinary Research* 51 (1): 27. doi:10.1186/s13567-020-00748-2.

Eichner, Miriam, Jonas Protze, Anna Piontek, Gerd Krause, and Jörg Piontek. 2017. "Targeting and alteration of tight junctions by bacteria and their virulence factors such as Clostridium perfringens enterotoxin." *Pflügers Archiv: European Journal of Physiology* 469 (1): 77-90. doi:10.1007/S00424-016-1902-X.

Farzan, Abdolvahab, Jasmina Kircanski, Josepha DeLay, Glenn Soltes, J. Glenn Songer, Robert Friendship, and John F. Prescott. 2013. "An investigation into the association between cpb2-encoding Clostridium perfringens type A and diarrhea in neonatal piglets." *Canadian Journal of Veterinary Research-revue Canadienne De Recherche Veterinaire* 77 (1): 45-53.

Fennessey, Christine M., Jinsong Sheng, Donald H. Rubin, and Mark S. McClain. 2012. "Oligomerization of Clostridium perfringens Epsilon Toxin Is Dependent upon Caveolins 1 and 2." *PLOS ONE* 7 (10). doi:10.1371/JOURNAL.PONE.0046866.

Fernandez-Miyakawa, Mariano E., B. Helen Jost, Stephen J. Billington, and Francisco A. Uzal. 2008. "Lethal effects of Clostridium perfringens epsilon toxin are potentiated by alpha and

perfringolysin-O toxins in a mouse model." *Veterinary Microbiology* 127 (3): 379-385. doi:10.1016/J.VETMIC.2007.09.013.

Freedman, John C., Bruce A. McClane, and Francisco A. Uzal. 2016. "New insights into Clostridium perfringens epsilon toxin activation and action on the brain during enterotoxemia." *Anaerobe* 41: 27-31. doi:10.1016/j.anaerobe.2016.06.006.

Giannitti, F., M. Macias Rioseco, J. P. García, J. Beingesser, Leslie Woods, Birgit Puschner, and Francisco A Uzal. 2014. "Diagnostic Exercise: Hemolysis and Sudden Death in Lambs." *Veterinary Pathology* 51 (3): 624-627. https://academic.microsoft.com/paper/2099399774.

Gibert, Maryse, Colette Jolivet-Renaud, and Michel R. Popoff. 1997. "Beta2 toxin, a novel toxin produced by Clostridium perfringens." *Gene* 203 (1): 65-73. doi:10.1016/S0378-1119(97)00493-9.

Gibert, Maryse, Jean Christophe Marvaud, Yannick Pereira, Martha L. Hale, Bradley G. Stiles, Patrice Boquet, Christophe Lamaze, and Michel R. Popoff. 2007. "Differential requirement for the translocation of clostridial binary toxins: Iota toxin requires a membrane potential gradient." *FEBS Letters* 581 (7): 1287-1296. doi:10.1016/J.FEBSLET.2007.02.041.

Gohari, Iman Mehdizadeh, Valeria R. Parreira, Victoria J. Nowell, Vivian M. Nicholson, Kaitlyn Oliphant, and John F. Prescott. 2015. "A novel pore-forming toxin in type A Clostridium perfringens is associated with both fatal canine hemorrhagic gastroenteritis and fatal foal necrotizing enterocolitis." *PLOS ONE* 10 (4). doi:10.1371/JOURNAL.PONE.0122684.

Goossens, Evy, Bonnie R. Valgaeren, Bart Pardon, Freddy Haesebrouck, Richard Ducatelle, Piet R. Deprez, and Filip Van Immerseel. 2017. "Rethinking the role of alpha toxin in Clostridium perfringens-associated enteric diseases : a review on bovine necro-haemorrhagic enteritis." *Veterinary Research* 48 (1): 9. doi:10.1186/S13567-017-0413-X.

Goossens, Evy, Stefanie Verherstraeten, Bonnie R. Valgaeren, Bart Pardon, Leen Timbermont, Stijn Schauvliege, Diego Rodrigo-Mocholí, et al. 2016. "The C-terminal domain of Clostridium perfringens alpha toxin as a vaccine candidate against bovine necrohemorrhagic enteritis." *Veterinary Research* 47 (1): 52-52. https://academic.microsoft.com/paper/2344934137.

Hu, Wen-Si, Hun Kim, and Ok Kyung Koo. 2018. "Molecular genotyping, biofilm formation and antibiotic resistance of enterotoxigenic Clostridium perfringens isolated from meat supplied to school cafeterias in South Korea." *Anaerobe* 52: 115-121. doi:10.1016/J.ANAEROBE.2018.06.011.

Hunter, S.E.C., J.E. Brown, P.C.F. Oyston, J. Sakurai, and R.W. Titball. 1993. "Molecular genetic analysis of beta-toxin of Clostridium perfringens reveals sequence homology with alpha-toxin, gamma-toxin, and leukocidin of Staphylococcus aureus." *Infection and Immunity* 61 (9): 3958-3965. doi:10.1128/IAI.61.9.3958-3965.1993.

Ivie, Susan E., and Mark S. McClain. 2012. "Identification of Amino Acids Important for Binding of Clostridium perfringens Epsilon Toxin to Host Cells and to HAVCR1." *Biochemistry* 51 (38): 7588-7595. doi:10.1021/BI300690A.

Jin, Fu, Osamu Matsushita, Sei-Ichi Katayama, Shengyong Jin, Chieko Matsushita, Junzaburo Minami, and Andakinobu Okabe. 1996. "Purification, characterization, and primary structure of Clostridium perfringens lambda-toxin, a thermolysin-like metalloprotease." *Infection and Immunity* 64 (1): 230-237. doi:10.1128/IAI.64.1.230-237.1996.

Jones, A. L., M. P. Dagleish, and George Caldow. 2015. "Clostridium perfringens type-D enterotoxaemia in cattle: the diagnostic significance of intestinal epsilon toxin." *Veterinary Record* 177 (15): 390-390. doi:10.1136/VR.103097.

Kalelkar, Sandeep, John Glushka, Herman vanHalbeek, Laura C. Morris, and Robert Cherniak. 1997. "Structure of the capsular polysaccharide of Clostridium perfringens Hobbs 5 as determined by NMR spectroscopy." *Carbohydrate Research* 299 (3): 119-128. doi:10.1016/S0008-6215(97)00010-4.

Karp, Peter D., Suzanne M. Paley, Peter E. Midford, Markus Krummenacker, Richard Billington, Anamika Kothari, Wai Kit Ong, Pallavi Subhraveti, Ingrid M Keseler, and Ron Caspi. 2021. "Pathway Tools version 23.0: Integrated Software for Pathway/Genome Informatics and Systems Biology." *arXiv* 1-111. doi:10.48550/arXiv.1510.03964.

Katoh, Kazutaka, Kazuharu Misawa, Kei-ichi Kuma, and Takashi Miyata. 2002. "MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform." *Nucleic Acids Research* 30 (14): 3059-3066. doi:10.1093/nar/gkf436.

Keokilwe, L., A. Olivier, W.P. Burger, H. Joubert, Estelle Hildegard Venter, and D. Morar-Leather. 2015. "Bacterial enteritis in ostrich (Struthio Camelus) chicks in the Western Cape Province, South Africa." *Poultry Science* 94 (6): 1177-1183. doi:10.3382/PS/PEV084.

Keyburn, Anthony L, John D Boyce, Paola Vaz, Trudi L Bannam, Mark E Ford, Dane Parker, Antonio Di Rubbo, Julian I Rood, and Robert J Moore. 2008. "NetB, a New Toxin That Is Associated with Avian Necrotic Enteritis Caused by Clostridium perfringens." *PLOS Pathogens* 4 (2): 26-26. doi:10.1371/JOURNAL.PPAT.0040026.

Kim, Ha-Young, Jae-Won Byun, In-Soon Roh, You-Chan Bae, Myoung-Heon Lee, Bumseok Kim, J. Glenn Songer, and Byeong Yeal Jung. 2013. "First isolation of Clostridium perfringens type E from a goat with diarrhea." *Anaerobe* 22: 141-143. doi:10.1016/J.ANAEROBE.2013.06.009.

Kiu, Raymond, and Lindsay J. Hall. 2018. "An update on the human and animal enteric pathogen Clostridium perfringens." *Emerging microbes & infections* 7 (1): 141-141. doi:10.1038/s41426-018-0144-8.

Kiu, Raymond, Shabhonam Caim, Sarah Alexander, Purnima Pachori, and Lindsay J. Hall. 2017. "Probing Genomic Aspects of the Multi-Host Pathogen Clostridium perfringens Reveals Significant Pangenome Diversity, and a Diverse Array of Virulence Factors." *Frontiers in Microbiology* 8: 2485-2485. https://academic.microsoft.com/paper/2771810836.

Krzywinski, Martin, Jacqueline Schein, Inac Birol, Joseph Connors, Randy Gascoyne, Doug Horsman, Steven J. Jones, and Marco A. Marra. 2009. "Circos: An information aesthetic for comparative genomics." *Genome Research* 19 (9): 1639-1645. doi:10.1101/gr.092759.109.

Lacey, Jake A., Theodore R. Allnutt, Ben Vezina, Thi Thu Hao Van, Thomas Stent, Xiaoyan Han, Julian I. Rood, et al. 2018. "Whole genome analysis reveals the diversity and evolutionary relationships between necrotic enteritis-causing strains of Clostridium perfringens." *BMC Genomics* 19 (1): 379-379. https://academic.microsoft.com/paper/2806209637.

LaMont, J.Thomas, Emily B. Sonnenblick, and Sara Rothman. 1979. "Role of clostridial toxin in the pathogenesis of clindamycin colitis in rabbits." *Gastroenterology* 76 (2): 356-361. doi:10.1016/0016-5085(79)90346-9.

Leipig-Rudolph, Miriam, Kathrin Busch, John F. Prescott, Iman Mehdizadeh Gohari, Christian M. Leutenegger, Walter Hermanns, Georg Wolf, Katrin Hartmann, Jutta Verspohl, and Stefan Unterer. 2018. "Intestinal lesions in dogs with acute hemorrhagic diarrhea syndrome associated with netF-positive Clostridium perfringens type A:." *Journal of Veterinary Diagnostic Investigation* 30 (4): 495-503. doi:10.1177/1040638718766983.

Li, Jihong, Sameera Sayeed, Susan Robertson, Jianming Chen, and Bruce A. McClane. 2011. "Sialidases Affect the Host Cell Adherence and Epsilon Toxin-Induced Cytotoxicity of Clostridium perfringens Type D Strain CN3718." *PLOS Pathogens* 7 (12). doi:10.1371/JOURNAL.PPAT.1002429.

Li, Jihong, Vicki Adams, Trudi Leanne Bannam, Kazuaki Miyamoto, Jorge P Garcia, Francisco A Uzal, Julian Ian Rood, and Bruce A McClane. 2013. "Toxin Plasmids of Clostridium perfringens." *Microbiology and Molecular Biology Reviews* 77 (2): 208-233. doi:10.1128/MMBR.00062-12.

Long, Jinzhao, Yake Xu, Liuyang Ou, Haiyan Yang, Yuanlin Xi, Shuaiyin Chen, and Guangcai Duan. 2019. "Diversity of CRISPR/Cas system in Clostridium perfringens." *Molecular Genetics and Genomics* 294 (5): 1263-1275. https://academic.microsoft.com/paper/2947415403.

Ma, Menglin, Jorge Vidal, Juliann Saputo, Bruce A. McClane, and Francisco A Uzal. 2011. "The VirS/VirR Two-Component System Regulates the Anaerobic Cytotoxicity, Intestinal Pathogenicity, and Enterotoxemic Lethality of Clostridium perfringens Type C Isolate CN3685." *Mbio* 2 (1). doi:10.1128/MBIO.00338-10.

McEwen, A.D., and R.S. Roberts. 1931. "“Struck”: Enteritis and Peritonitis of Sheep Caused by a Bacterial Toxin Derived from the Alimentary Canal." *Journal of Comparative Pathology* 44: 26-49. doi:10.1016/S0368-1742(31)80002-5.

Miclard, J., M. Jäggi, E. Sutter, M. Wyder, B. Grabscheid, and H. Posthaus. 2009. "Clostridium perfringens beta-toxin targets endothelial cells in necrotizing enteritis in piglets." *Veterinary Microbiology* 137 (3): 320-325. doi:10.1016/J.VETMIC.2009.01.025.

Minami, Junzaburo, Seiichi Katayama, Osamu Matsushita, Chieko Matsushita, and Akinobu Okabe. 1997. "Lambda-Toxin of Clostridium perfringens Activates the Precursor of Epsilon-Toxin by Releasing Its N- and C-Terminal Peptides." *Microbiology and Immunology* 41 (7): 527-535. doi:10.1111/j.1348-0421.1997.tb01888.x .

Miyakawa, Mariano E. Fernandez, Julian Saputo, Judy St. Leger, Birgit Puschner, Derek J. Fisher, Bruce A. McClane, and Francisco A. Uzal. 2007. "Necrotizing enterocolitis and death in a goat kid associated with enterotoxin (CPE)-producing Clostridium perfringens type A."

Canadian Veterinary Journal-revue Veterinaire Canadienne 48 (12): 1266-1269. https://academic.microsoft.com/paper/232587671.

Nagahama, Masahiro, Akiko Ohkubo, Masataka Oda, Keiko Kobayashi, Katsuhiko Amimoto, Kazuaki Miyamoto, and Jun Sakurai. 2011. "Clostridium perfringens TpeL Glycosylates the Rac and Ras Subfamily Proteins." *Infection and Immunity* 79 (2): 905-910. doi:10.1128/IAI.01019-10.

Nagahama, Masahiro, Akiwo Yamaguchi, Tohko Hagiyama, Noriko Ohkubo, Keiko Kobayashi, and Jun Sakurai. 2004. "Binding and Internalization of Clostridium perfringens Iota-Toxin in Lipid Rafts." *Infection and Immunity* 72 (6): 3267-3275. doi:10.1128/IAI.72.6.3267-3275.2004.

Nagahama, Masahiro, Hideki Hara, Mariano Fernandez-Miyakawa, Yukari Itohayashi, and Jun Sakurai. 2006. "Oligomerization of Clostridium perfringens epsilon-toxin is dependent upon membrane fluidity in liposomes." *Biochemistry* 45 (1): 296-302. doi:doi/10.1021/bi051805s.

NCBI. 1988. *National Center for Biotechnology Information (NCBI).* Accessed 2021. https://www.ncbi.nlm.nih.gov/.

Neeson, Brendan N., Graeme C. Clark, Helen S. Atkins, Bryan Lingard, and Richard W. Titball. 2007. "Analysis of protection afforded by a Clostridium perfringens α-toxoid against heterologous clostridial phospholipases C." *Microbial Pathogenesis* 43 (4): 161-165. doi:10.1016/J.MICPATH.2007.05.004.

Nestorovich, Ekaterina M., Vladimir A. Karginov, and Sergey M. Bezrukov. 2010. "Polymer Partitioning and Ion Selectivity Suggest Asymmetrical Shape for the Membrane Pore Formed by Epsilon Toxin." *Biophysical Journal* 99 (3): 782-789. doi:10.1016/j.bpj.2010.05.014.

Ngamwongsatit, Bhinyada, Wimonrat Tanomsridachchai, Orasa Suthienkul, Supanee Urairong, Wichian Navasakuljinda, and Tavan Janvilisri. 2016. "Multidrug resistance in Clostridium perfringens isolated from diarrheal neonatal piglets in Thailand." *Anaerobe* 38: 88-93. doi:10.1016/J.ANAEROBE.2015.12.012.

Oda, Masataka, Ryota Shiihara, Yuka Ohmae, Michiko Kabura, Teruhisa Takagishi, Keiko Kobayashi, Masahiro Nagahama, et al. 2012. "Clostridium perfringens alpha-toxin induces the release of IL-8 through a dual pathway via TrkA in A549 cells." *Biochimica et Biophysica Acta* 1822 (10): 1581-1589. doi:10.1016/J.BBADIS.2012.06.007.

Oda, Masataka, Yutaka Terao, Jun Sakurai, and Masahiro Nagahama. 2015. "Membrane-Binding Mechanism of Clostridium perfringens Alpha-Toxin." *Toxins* 7 (12): 5268-5275. doi:10.3390/TOXINS7124880.

Ohtani, Kaori, Hameem I. Kawsar, Kayo Okumura, Hideo Hayashi, and Tohru Shimizu. 2003. "The VirR/VirS regulatory cascade affects transcription of plasmid-encoded putative virulence genes in Clostridium perfringens strain 13." *Fems Microbiology Letters* 222 (1): 137-141. doi:10.1016/S0378-1097(03)00255-6.

Ohtani, Kaori, Hideki Hirakawa, Kousuke Tashiro, Satoko Yoshizawa, Satoru Kuhara, and Tohru Shimizu. 2010. "Identification of a two-component VirR/VirS regulon in Clostridium perfringens." *Anaerobe* 16 (3): 258-264. doi:10.1016/J.ANAEROBE.2009.10.003.

Ohtani, Kaori, Yonghui Yuan, Sufi Hassan, Ruoyu Wang, Yun Wang, and Tohru Shimizu. 2009. "Virulence Gene Regulation by the agr System in Clostridium perfringens." *Journal of Bacteriology* 191 (12): 3919-3927. doi:10.1128/JB.01455-08.

Okumura, Kayo, Kaori Ohtani, Hideo Hayashi, and Tohru Shimizu. 2008. "Characterization of Genes Regulated Directly by the VirR/VirS System in Clostridium perfringens." *Journal of Bacteriology* 190 (23): 7719-7727. doi:10.1128/JB.01573-07.

Paley, Suzanne, O'Maille Paul, Daniel Weaver, and Peter D Karp. 2016. "Pathway collages: personalized multi-pathway diagrams." *BMC Bioinformatics* 529. doi:10.1186/s12859-016-1382-1.

Papatheodorou, Panagiotis, Jan E. Carette, George W. Bell, Carsten Schwan, Gregor Guttenberg, Thijn R. Brummelkamp, and Klaus Aktories. 2011. "Lipolysis-stimulated lipoprotein receptor (LSR) is the host receptor for the binary toxin Clostridium difficile transferase (CDT)." *Proceedings of the National Academy of Sciences of the United States of America* 108 (39): 16422-16427. doi:10.1073/PNAS.1109772108.

Park, Miseon, Alejandro P. Rooney, David W. Hecht, Jihong Li, Bruce A. McClane, Rajesh Nayak, Donald D. Paine, and Fatemeh Rafii. 2010. "Phenotypic and genotypic characterization of tetracycline and minocycline resistance in Clostridium perfringens." *Archives of Microbiology* 192 (10): 803-810. doi:10.1007/S00203-010-0605-5.

Parks, Donovan H., Michael Imelfort, Connor T. Skennerton, Philip Hugenholtz, and Gene W. Tyson. 2015. "CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes." *Genome Research* 25 (7): 1043-1055. doi:10.1101/gr.186072.114.

Parrello, Bruce, Rory Butler, Philippe Chlenski, Robert Olson, Jamie Overbeek, Gordon Pusch, Veronika Vonstein, and Ross Overbeek. 2019. "A machine learning-based service for estimating quality of genomes using PATRIC." *BMC Bioinformatics.* doi:10.1186/s12859-019-3068-y.

Popoff, Michel R. 2011. "Epsilon toxin: a fascinating pore-forming toxin." *FEBS Journal* 278 (23): 4602-4615. doi:10.1111/j.1742-4658.2011.08145.x.

Redondo, L.M., M. Farber, A. Venzano, B.H. Jost, Y.R. Parma, and M.E. Fernandez-Miyakawa. 2013. "Sudden death syndrome in adult cows associated with Clostridium perfringens type E." *Anaerobe* 20: 1-4. doi:10.1016/j.anaerobe.2013.01.001.

Redondo, Leandro M., Enzo A. Redondo, Gabriela C. Dailoff, Carlos L. Leiva, Juan M. Díaz-Carrasco, Octavio A. Bruzzone, Adriana Cangelosi, Patricia Geoghegan, and Mariano E. Fernandez-Miyakawa. 2017. "Effects of Clostridium perfringens iota toxin in the small intestine of mice." *Anaerobe* 48: 83-88. doi:10.1016/j.anaerobe.2017.07.007.

Rehman, H., W. A. Awad, I. Lindner, M. Hess, and J. Zentek. 2006. "Clostridium perfringens alpha toxin affects electrophysiological properties of isolated jejunal mucosa of laying hens." *Poultry Science* 85 (7): 1298-1302. doi:10.1093/ps/85.7.1298.

Robertson, Susan L., James G. Smedley, Usha Singh, Ganes Chakrabarti, Christina M. Van Itallie, James M. Anderson, and Bruce A. McClane. 2007. "Compositional and stoichiometric analysis of Clostridium perfringens enterotoxin complexes in Caco-2 cells and claudin 4 fibroblast transfectants." *Cellular Microbiology* 9 (11): 2734-2755. doi:10.1111/J.1462-5822.2007.00994.X.

Rodgers, Katherine, Cindy Grove Arvidson, and Stephen Melville. 2011. "Expression of a Clostridium perfringens Type IV Pilin by Neisseria gonorrhoeae Mediates Adherence to Muscle Cells." *Infection and Immunity* 79 (8): 3096-3105. doi:10.1128/IAI.00909-10.

Rood, Julian I., Anthony L. Keyburn, and Robert J. Moore. 2016. "NetB and necrotic enteritis: the hole movable story." *Avian Pathology* 45 (3): 295-301. doi:10.1080/03079457.2016.1158781.

Rood, Julian I., Vicki Adams, Jake Lacey, Dena Lyras, Bruce A. McClane, Stephen B. Melville, Robert J. Moore, et al. 2018. "Expansion of the Clostridium perfringens toxin-based typing scheme." *Anaerobe* 53: 5-10. doi:10.1016/j.anaerobe.2018.04.011.

Sakurai, Jun, Masahiro Nagahama, and Masataka Oda. 2004. "Clostridium perfringens Alpha-Toxin: Characterization and Mode of Action." *Journal of Biochemistry* 136 (5): 569-574. https://academic.microsoft.com/paper/2029626396.

Sayeed, Sameera, Francisco A Uzal, Derek J Fisher, Julian Saputo, Jorge E Vidal, Yue Chen, Phalguni Gupta, Julian Ian Rood, and Bruce A McClane. 2007. "Beta toxin is essential for the intestinal virulence of Clostridium perfringens type C disease isolate CN3685 in a rabbit ileal loop model." *Molecular Microbiology* 67 (1): 15-30. doi:10.1111/j.1365-2958.2007.06007.

Shimizu, Takeshi, Kensuke Shima, Ken-ichi Yoshino, Kazuyoshi Yonezawa, Tohru Shimizu, and Hideo Hayashi. 2002. "Proteome and Transcriptome Analysis of the Virulence Genes Regulated by the VirR/VirS System in Clostridium perfringens." *Journal of Bacteriology* 184 (10): 2587-2594. doi:10.1128/JB.184.10.2587-2594.2002.

Shimizu, Tohru, Kaori Ohtani, Hideki Hirakawa, Kenshiro Ohshima, Atsushi Yamashita, Tadayoshi Shiba, Naotake Ogasawara, Masahira Hattori, Satoru Kuhara, and Hideo Hayashi. 2002. "Complete genome sequence of Clostridium perfringens, an anaerobic flesh-eater." *Proceedings of the National Academy of Sciences of the United States of America* 99 (2): 996-1001. https://academic.microsoft.com/paper/2146755217.

Shrestha, Archana, Iman Mehdizadeh Gohari, and Bruce A. McClane. 2019. "RIP1, RIP3, and MLKL Contribute to Cell Death Caused by Clostridium perfringens Enterotoxin." *Mbio* 10 (6). doi:10.1128/MBIO.02985-19.

Smedley, James G., Juliann Saputo, Jacquelyn C. Parker, Mariano E. Fernandez-Miyakawa, Susan L. Robertson, Bruce A. McClane, and Francisco A. Uzal. 2008. "Noncytotoxic Clostridium perfringens Enterotoxin (CPE) Variants Localize CPE Intestinal Binding and Demonstrate a

Relationship between CPE-Induced Cytotoxicity and Enterotoxicity." *Infection and Immunity* 76 (8): 3793-3800. doi:10.1128/IAI.00460-08.

Stamatakis, Alexandros. 2014. "RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies." *Bioinformatics* 30 (9): 1312-1313. doi:10.1093/bioinformatics/btu033.

Stokka, Gl, Aj Edwards, Mf Spire, Rt Brandt, and Je Smith. 1994. "INFLAMMATORY RESPONSE TO CLOSTRIDIAL VACCINES IN FEEDLOT CATTLE." *Javma-journal of The American Veterinary Medical Association* 204 (3): 415-419.

Takehara, Masaya, Teruhisa Takagishi, Soshi Seike, Kaori Ohtani, Keiko Kobayashi, Kazuaki Miyamoto, Tohru Shimizu, and Masahiro Nagahama. 2016. "Clostridium perfringens α-Toxin Impairs Innate Immunity via Inhibition of Neutrophil Differentiation." *Scientific Reports* 6 (1): 28192-28192. doi:10.1038/SREP28192.

Tamai, Eiji, Tetsuya Ishida, Shigeru Miyata, Osamu Matsushita, Hirofumi Suda, Shoji Kobayashi, Hiroshi Sonobe, and Akinobu Okabe. 2003. "Accumulation of Clostridium perfringens Epsilon-Toxin in the Mouse Kidney and Its Possible Biological Significance." *Infection and Immunity* 71 (9): 5371-5375. doi:10.1128/IAI.71.9.5371-5375.2003.

Tatusova, Tatiana, Michael DiCuccio, Azat Badretdin, Vyacheslav Chetvernin, Eric P Nawrocki, Leonid Zaslavsky, Alexandre Lomsadze, Kim D Pruitt, Mark Borodovsky, and James Ostell. 2016. "NCBI prokaryotic genome annotation pipeline." *Nucleic Acids Res.* 44 (14): 6614-6624. doi:10.1093/nar/gkw569.

Uzal, Francisco A, J. P. Wong, W. R. Kelly, and J. Priest. 1999. "Antibody response in goats vaccinated with liposome-adjuvanted Clostridium perfringens type D epsilon toxoid." *Veterinary Research Communications* 23 (3): 143-150. doi:10.1023/A:1006206216220.

Uzal, Francisco A, W. R. Kelly, W. E. Morris, J. Bermudez, and M. Baisón. 2004. "The Pathology of Peracute Experimental Clostridium Perfringens Type D Enterotoxemia in Sheep." *Journal of Veterinary Diagnostic Investigation* 16 (5): 403-411. doi:10.1177/104063870401600506.

Uzal, Francisco A., and J. Glenn Songer. 2008. "Diagnosis of Clostridium Perfringens Intestinal Infections in Sheep and Goats." *Journal of Veterinary Diagnostic Investigation* 20 (3): 253-265. doi:10.1177/104063870802000301.

Vandekerckhove, Joël, Beate Schering, Michael Bärmann, and Klaus Aktories. 1987. "Clostridium perfringens iota toxin ADP-ribosylates skeletal muscle actin in Arg-177." *FEBS Letters* 225: 48-52. doi:10.1016/0014-5793(87)81129-8.

Varga, John J., Blair Therit, and Stephen B. Melville. 2008. "Type IV Pili and the CcpA Protein Are Needed for Maximal Biofilm Formation by the Gram-Positive Anaerobic Pathogen Clostridium perfringens." *Infection and Immunity* 76 (11): 4944-4951. doi:10.1128/IAI.00692-08.

Varga, John J., Van Nguyen, David K. O'Brien, Katherine Rodgers, Richard A. Walker, and Stephen B. Melville. 2006. "Type IV pili-dependent gliding motility in the Gram-positive pathogen

Clostridium perfringens and other Clostridia." *Molecular Microbiology* 62 (3): 680-694. doi:10.1111/J.1365-2958.2006.05414.X.

Verherstraeten, Stefanie, Evy Goossens, Bonnie Valgaeren, Bart Pardon, Leen Timbermont, Freddy Haesebrouck, Richard Ducatelle, et al. 2015. "Perfringolysin O: The Underrated Clostridium perfringens Toxin?" *Toxins* 7 (5): 1702-1721. doi:10.3390/TOXINS7051702.

Verherstraeten, Stefanie, Evy Goossens, Bonnie Valgaeren, Bart Pardon, Leen Timbermont, Karen Vermeulen, Stijn Schauvliege, et al. 2013. "The synergistic necrohemorrhagic action of Clostridium perfringens perfringolysin and alpha toxin in the bovine intestine and against bovine endothelial cells." *Veterinary Research* 44 (1): 45-45. doi:10.1186/1297-9716-44-45.

Vos, P, G Garrity, D Jones, N R Krieg, W Ludwig, F A Rainey, K H Schleifer, and W B Whitman. 2005. *Bergey's Manual of Systematic Bacteriology.* 2. Vol. 2. New York, NY: Springer-Verlag.

Wade, Ben, Anthony L Keyburn, Torsten Seemann, Julian I Rood, and Robert J Moore. 2015. "Binding of Clostridium perfringens to collagen correlates with the ability to cause necrotic enteritis in chickens." *Veterinary Microbiology* 180 (3): 299-303. doi:10.1016/J.VETMIC.2015.09.019.

Wagley, Sariqa, Monika Bokori-Brown, Helen Morcrette, Andrea Malaspina, Caroline D'Arcy, Sharmilee Gnanapavan, Nicholas Lewis, et al. 2019. "Evidence of Clostridium perfringens epsilon toxin associated with multiple sclerosis." *Multiple Sclerosis Journal* 25 (5): 653-660. doi:10.1177/1352458518767327.

Wattam, Alice R, David Abraham, Oral Dalay, Terry L Disz, Timothy Driscoll, Joseph L Gabbard, Joseph J Gillespie, et al. 2014. "PATRIC, the bacterial bioinformatics database and analysis resource." *Nucleic Acids Research* 42 (D1): D581-D591. doi:10.1093/nar/gkt1099.

Williamson, E.D., and R.W. Titball. 1993. "A genetically engineered vaccine against the alpha-toxin of Clostridium perfringens protects mice against experimental gas gangrene." *Vaccine* 11 (12): 1253-1258. doi:10.1016/0264-410X(93)90051-X.

Yan, Xu-Xia, Corrine J. Porter, Simon P. Hardy, David Steer, A. Ian Smith, Noelene S. Quinsey, Victoria Hughes, et al. 2013. "Structural and Functional Analysis of the Pore-Forming Toxin NetB from Clostridium perfringens." *Mbio* 4 (1): 1-9. doi:10.1128/MBIO.00019-13.

Yonogi, Shinya, Shigeaki Matsuda, Takao Kawai, Tomoko Yoda, Tetsuya Harada, Yuko Kumeda, Kazuyoshi Gotoh, et al. 2014. "BEC, a Novel Enterotoxin of Clostridium perfringens Found in Human Clinical Isolates from Acute Gastroenteritis Outbreaks." *Infection and Immunity* 82 (6): 2390-2399. doi:10.1128/IAI.01759-14.

Yu, Qiang, Dion Lepp, Iman Mehdizadeh Gohari, Tao Wu, Hongzhuan Zhou, Xianhua Yin, Hai Yu, et al. 2017. "The Agr-Like Quorum Sensing System Is Required for Pathogenesis of Necrotic Enteritis Caused by Clostridium perfringens in Poultry." *Infection and Immunity* 85 (6). doi:10.1128/IAI.00975-16.

APPENDIX A: GENOME STATISTICS

Table A1: Isolation data of selected strains

STRAIN NAME	ACCESSION NUMBER	COUNTRY OF ISOLATION	YEAR OF ISOLATION
ATCC13124	CP000246.1	N/A	2005
STRAIN 37	PJSN01000000.1	Belgium	2005
STRAIN 67	PJSM01000000.1	Belgium	2005
STRAIN 98.787 18-2	PJTE01000000.1	Denmark	2009
DEL1	CP019576.1	USA	2002
EHE-NE18	CP025501.1	Australia	2002
EHE-NE7	PJTC01000000.1	Australia	2002
EUR-NE15	PJTB01000000.1	Australia	1995
FC2	PJTA01000000.1	USA	1996
GNP-1	PJSZ01000000.1	USA	1996
ITX1105-12MP	PJSY01000000.1	USA	1993
PENNINGTON	PJSQ01000000.1	USA	2007
SOM-NE35	PJTM01000000.1	Australia	2011
TAM-NE38	PJTL01000000.1	Australia	2012
TAM-NE40	PJSK01000000.1	Australia	2013
TAM-NE46	PJSI01000000.1	Australia	1995
UDE 95-1372	PJTI01000000.1	USA	1993
WARREN	PJTH01000000.1	Australia	2010
WER-NE36	PJTG01000000.1	USA	2005
C. VENTRICULI STRAIN 17	BCMW00000000.1	Japan	2005

Table A2: Sequencing data of selected strains

STRAIN NAME	SEQUENCING PLATFORM	SEQUENCING DEPTH	GENOME SIZE (MB)
ATCC13124	Illumina MiSeq	125x	3.26
STRAIN 37	Illumina MiSeq	122x	3.48
STRAIN 67	Illumina MiSeq	117x	3.45
STRAIN 98.787 18-2	Illumina HiSeq	300.0x	3.61
DEL1	PacBio	393.0x	3.81
EHE-NE18	Illumina MiSeq	34x	3.66
EHE-NE7	Illumina MiSeq	63x	3.43
EUR-NE15	Illumina MiSeq	28x	3.58
FC2	Illumina MiSeq	30x	3.73
GNP-1	Illumina MiSeq	95x	3.73
ITX1105-12MP	Illumina MiSeq	33x	3.65
PENNINGTON	Illumina MiSeq	167x	3.75
SOM-NE35	454	29x	3.57
TAM-NE38	454	24x	3.65
TAM-NE40	Illumina MiSeq	30x	3.65
TAM-NE46	Illumina MiSeq	38x	3.68
UDE 95-1372	Illumina MiSeq	121x	3.64
WARREN	Illumina MiSeq	12x	3.41
WER-NE36	Illumina MiSeq	125x	3.78

Table A3: Genomic data of selected strains

STRAIN NAME	COARSE CONSISTENCY	FINE CONSISTENCY	CHECKM COMPLETENESS	CHECKM CONTAMINATION
ATCC13124	99.9	99.9	100	0
STRAIN 37	99.8	98.7	100	0
STRAIN 67	99.8	99	100	0
STRAIN 98.787 18-2	99.9	99.2	100	0
DEL1	99.9	98.5	100	0
EHE-NE18	99.9	99.2	100	0
EHE-NE7	99.8	99.1	100	0
EUR-NE15	99.9	99.2	100	0
FC2	99.9	99.2	100	0
GNP-1	99.8	99.2	100	0
ITX1105-12MP	99.9	99.2	100	0
PENNINGTON	99.9	99.3	100	0
SOM-NE35	99.8	99	100	1.1
TAM-NE38	99.9	98.8	100	0
TAM-NE40	99.9	98.8	100	0
TAM-NE46	99.9	98.7	100	2.2
UDE 95-1372	99.8	99	100	0
WARREN	99.7	98.9	100	0
WER-NE36	99.8	98.4	100	1.1

Table A4: Annotation data of selected strains

STRAIN NAME	PATRIC CDS	REFSEQ CDS
ATCC13124	2866	2876
STRAIN 37	3229	3101
STRAIN 67	3211	3079
STRAIN 98.787 18-2	3433	3283
DEL1	3679	3435
EHE-NE18	3398	3253
EHE-NE7	3199	3059
EUR-NE15	3377	3242
FC2	3272	3145
GNP-1	3621	3451
ITX1105-12MP	3518	3364
PENNINGTON	3655	3486
SOM-NE35	3339	3201
TAM-NE38	3527	3355
TAM-NE40	3561	3369
TAM-NE46	3655	3488
UDE 95-1372	3515	3357
WARREN	3648	3489
WER-NE36	3279	3088

APPENDIX B: PHYLOGENETIC RECONSTRUCTION

Tree Analysis Statistics

Requested genomes	20
Genomes with data	20
Max allowed deletions	0
Max allowed duplications	0
Single-copy genes requested	1000
Single-copy genes found	739
Num protein alignments	739
Alignment program	mafft
Protein alignment time	839.5 seconds
Num aligned amino acids	243051
Num CDS alignments	739
Num aligned nucleotides	729153
Best protein model found by RAxML	CPREV
Branch support method	RAxML Fast Bootstrapping
RAxML likelihood	-2674615.0587
RAxML version	8.2.11
RAxML time	477.2 seconds
Total time	1586.4 seconds

RAxML Command Line

Goal: Analyze proteins with model 'AUTO' to find best substitution model.
raxmlHPC-PTHREADS-SSE3 -s Ventriculiout_proteins.phy -n Ventriculiout_proteins -m PROTCATAUTO -p 12345 -T 12 -e 10
Process time: 81.670 seconds

Goal: Find best tree.
raxmlHPC-PTHREADS-SSE3 -s Ventriculiout.phy -n Ventriculiout -m GTRCAT -q Ventriculiout.partitions -p 12345 -T 12 -f a -x 12345 -N 100
Process time: 395.573 seconds

RAxML Codon and Amino Acid Partitions

```
DNA, codon1 = 1-729153\3
```

```
DNA, codon2 = 2-729153\3
DNA, codon3 = 3-729153\3
CPREV, proteins = 729154-972204
```

Genome Statistics

GenomeId	Total Genes	Single Copy	Used	Name
1267.5	1956	739	739	Clostridium ventriculi strain strains 14 and 17
195103.10	2388	739	739	Clostridium perfringens ATCC 13124
1502.429	2492	739	739	Clostridium perfringens strain WER-NE36
1502.459	2493	739	739	Clostridium perfringens strain EHE-NE7
1502.438	2522	739	739	Clostridium perfringens strain 37
1502.436	2535	739	739	Clostridium perfringens strain 67
1502.446	2545	739	739	Clostridium perfringens strain FC2
1502.434	2589	739	739	Clostridium perfringens strain SOM-NE35
1502.449	2645	739	739	Clostridium perfringens strain 98.78718-2
1502.447	2650	739	739	Clostridium perfringens strain EUR-NE15
1502.460	2670	739	739	Clostridium perfringens strain EHE-NE18
1502.427	2691	739	739	Clostridium perfringens strain UDE 95-1372
1502.444	2696	739	739	Clostridium perfringens strain ITX1105-12MP
1502.433	2699	739	739	Clostridium perfringens strain TAM-NE40
1502.435	2700	739	739	Clostridium perfringens strain TAM-NE38
1502.430	2739	739	739	Clostridium perfringens strain TAM-NE46
1502.440	2755	739	739	Clostridium perfringens strain Pennington
1502.445	2756	739	739	Clostridium perfringens strain GNP-1
1502.451	2760	739	739	Clostridium perfringens strain Warren
1502.249	2801	739	739	Clostridium perfringens strain Del1

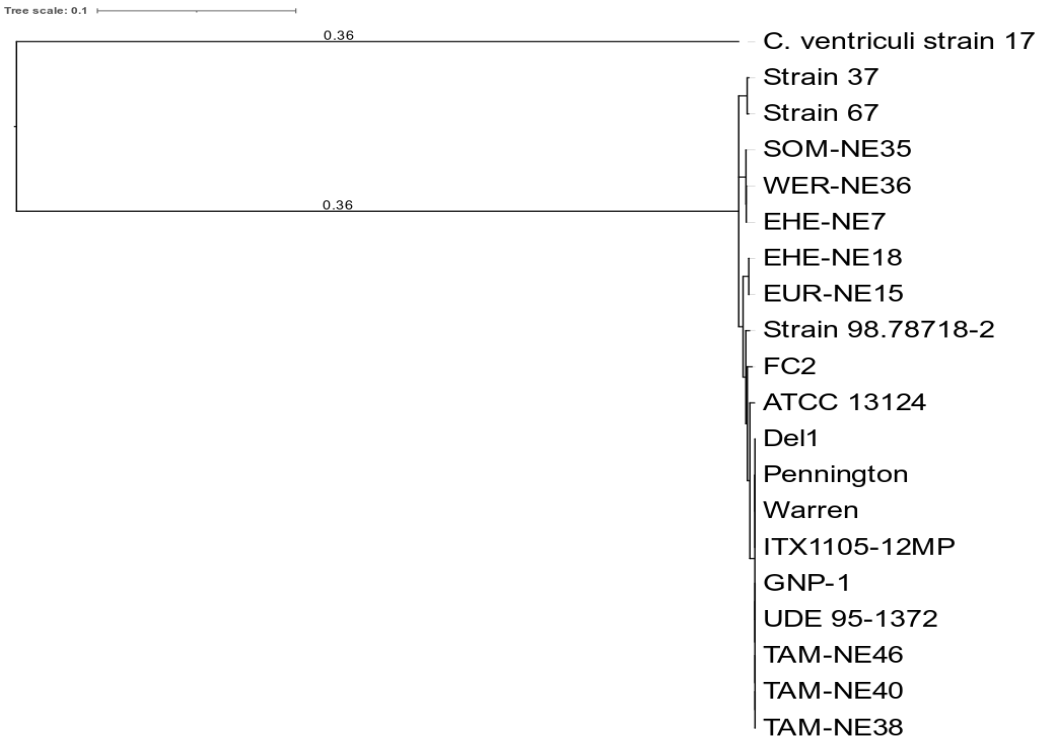

Figure B1: Phylogenetic reconstruction using *C. ventriculi* as outgroup

Phylogenetic reconstruction performed using the parameters in Appendix B. All nodes are supported with bootstrap values >80%. The root without the outgroup (used for Figure 1) was identified to be the node between the branch containing Strain 37, Strain 67 and SOM-NE35, WER-NE36, and EHE-NE7