Draft genome sequence data of *Clostridium perfringens* FA isolated from the faecal material of the critically endangered African wild dog, *Lycaon pictus*

Fatima Aguilar Sanchez\(^a\), Janelle Brandt\(^b\), Richard William McLaughlin\(^a, \ast\)

\(^a\) School of Liberal Arts and Sciences, Gateway Technical College, Kenosha, WI 53144, United States
\(^b\) Great Plains Zoo & Delbridge Museum of Natural History, Sioux Falls, SD 57104, United States

**A R T I C L E  I N F O**

**Article history:**
Received 1 August 2022  
Revised 15 October 2022  
Accepted 17 October 2022  
Available online 21 October 2022

**Dataset link:** Clostridium perfringens FA (Original data)

**Keywords:**  
Lycaon pictus  
African wild dog  
Faecal material  
Clostridium perfringens  
Whole genome sequencing

**A B S T R A C T**

In this present article the draft sequence data for *Clostridium perfringens* FA, which was isolated from the faecal material of a critically endangered African Wild dog, is reported. The bacterium is widely distributed in the environment and in the normal intestinal flora of humans and animals. The genome of strain *C. perfringens* FA was assembled into 21 contigs with a total length of 3,044,349 bp and a GC content of 28.20%. There are 2742 CDS, 70 tRNAs and 5 rRNAs. Five putative virulence genes were detected. There were no plasmid replicons found. The genome of few environmental isolates has been sequenced. The draft genome of strain FA can be compared to disease causing isolates cultured from humans to aid in a better understanding of the pathogenesis of the bacterium.

© 2022 The Author(s). Published by Elsevier Inc.  
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

\(^\ast\) Corresponding author.  
E-mail address: mclaughlin@gtc.edu (R.W. McLaughlin).

https://doi.org/10.1016/j.dib.2022.108685

2352-3409 © 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Specifications Table

| Subject                        | Biological Sciences                        |
|-------------------------------|--------------------------------------------|
| Specific subject area         | Environmental Microbiology; Bacterial genomics |
| Type of data                  | Genome sequence data                        |
| How the data were acquired    | Whole genome sequencing: NextSeq 550 (Illumina, San Diego, USA) |
|                               | De novo assembly: SPAdes version 3.15.4     |
|                               | Annotation: RAST tool kit (RASTtk) release 2018–0531 |
|                               | Phylogenetic analysis: Codon Tree           |
| Data format                   | Raw reads                                   |
| Description of data collection| For genomic sequencing a colony of the bacterium was grown anaerobically at 30 °C for 48 h on blood agar and then sent to the SeqCenter (https://www.seqcenter.com/). Sequencing was performed on a NextSeq 550 (Illumina, San Diego, USA). The genome was assembled using SPAdes version 3.13.0 and annotated using the RAST tool kit (RASTtk). |
| Data source location          | Strain FA was isolated from the faecal material of an African Wild dog living at the Great Plains Zoo & Delbridge Museum of Natural History, Sioux Falls SD, USA. |
| Data accessibility            | This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JANFDC0000000000. The version described in this paper is version JANFDC0000000000. Raw reads were deposited in SRA under accession no. SRX16308591 BioProject under accession no. PRJNA859381 BioSample under accession no. SAMN29783039 |
| Related research article      | None                                        |

Value of the Data

- A comprehensive analysis of the genome of strain FA can be done to determine virulence factors and the presence of other genes of interest.
- The genome of strain FA can be compared to the genomes of other *C. perfringens* strains which were isolated from human infections.
- A phylogenetic relationship can be established between closely related isolates using a core-genome-based MLST (cgMLST) approach.

1. Data Description

*C. perfringens* is part of the normal intestinal flora of humans and animals. The bacterium can also be a pathogen to both humans and animals. For example, type A can gas gangrene (myonecrosis) [1] and can be responsible for foodborne poisoning in industrialized countries [2]. This type of food-poisoning outbreaks primarily involves high protein foods of animal origin, such as meat, meat products and meat dishes [3].

The genome of strain FA was assembled from 2 × 151 bp reads raw reads into 21 contigs with a total length of 3, 044, 349 bp. A bar graph showing the percent each contig contributes to the total chromosome length is shown in Fig. 1. The contig L50 is 2 and the contig N50 is 501, 639. There are 2742 CDS, 2104 proteins with functional assignments, 638 hypothetical proteins, 70 tRNAs and 5 rRNAs. To determine the identity of the bacterium the average nucleotide index (ANI) value for our isolate was determined by comparing strain FA to the type strain *C. perfringens* ATCC 13124T. A value of 95.92% identity was obtained. Since the cut-off proposed for a species boundary is 95% [4] strain FA can be positively identified as *C. perfringens*. There were no plasmid replicons identified. A phylogenetic tree displaying the relationships between strain FA and other *Clostridium* species is shown in Fig. 2.

Using *in silico* methods putative virulence factors were identified. The bacterium was found to process a *cpa* gene (alpha toxin). However, the bacterium did not possess the *cpb* (beta toxin),
Fig. 1. A histogram depicting the percent contribution of each contig to the total chromosome length.

Fig. 2. A phylogenetic analysis created by Codon Tree using eleven *Clostridium* genomes. The genome sequenced in this study is in bold. Brackets indicate the GenBank accession number. Superscript T indicates a type strain.

cpb2 (β2-toxin), etx (epsilon toxin), iap (iota toxin), cpe (enterotoxin) or netB gene. As a result, this strain is type A [1]. In addition, strain FA has the perfringolysin O (pfoA) gene, which encodes a type of toxin produced by most *C. perfringens* strains [5]. Strain FA also contains a collagenase (colA), exo-alpha-sialidase (nanH) and the clostripain (cloSI) gene.
2. Experimental Design, Materials and Methods

A male African wild dog Monkey Wrench, living in the Great Plains Zoo & Delbridge Museum of Natural History, Sioux Falls SD, was used in this study. Faecal samples were transported to a − 40 °C freezer and stored until used for laboratory experiments. In total 10 mg of faecal material was added to 1 ml of 70% ethanol and incubated for 30 min with mixing by inversion every 10 min [6]. 100 μl aliquots of the suspension were then plated on sheep blood agar plates (Remel, Santa Fe CA, USA). The plates were incubated anaerobically at 30 °C for two days. Bacterial colonies were sub-cultured two separate times to ensure a pure culture was obtained.

A colony of *Clostridium perfringens* strain FA was sent to the SeqCenter (https://www.seqcenter.com/) for genome sequencing. A summary of the methods can be found in the paper by Baym et al. [7]. Sequencing was performed on a NextSeq 550 (Illumina, San Diego, USA). Default parameters were used for all software unless otherwise stated. The genome was assembled using SPAdes version 3.13.0 [8] and annotated using the RAST tool kit (RASTtk) release 2018–0531 [9]. Genome-based identification was then done by submitting the draft genome sequence to the ANI calculator [10]. A phylogenetic tree was constructed using Codon Tree [11] (Fig. 2). Plasmid replicons were predicted using PlasmidFinder 2.1 [12]. Finally, the BacWGSTdb database [13] was used to find the presence or absence of putative virulence factor genes in strain FA. The parameters single genome analysis, *Clostridioides difficile* as the species, 1000 as the SNP threshold, cgMLST as the MLST scheme and 1000 as the MLST threshold were used.

**Ethics Statements**

Approval was granted by the Animal Care management team from the Great Plains Zoo & Delbridge Museum of Natural History for the collection of faecal samples.

**CRediT Author Statement**

**Fatima Aguilar Sanchez:** Conceptualization, Methodology, Investigation; **Janelle Brandt:** Resources, Writing – review & editing; **Richard William McLaughlin:** Conceptualization, Methodology, Investigation, Data curation, Writing – original draft preparation, Writing – review & editing.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Data Availability**

*Clostridium perfringens* FA (Original data) (GenBank).

**Acknowledgments**

This research was done as part of a Provost Honors project under the leadership of Zina Haywood, Executive Vice President/Provost. We also thank Jorge Nieto, Nicole Dutton and Donald Zakutansky for their enthusiastic support of this research. This project was supported by funding provided by Gateway Technical College and by the Gateway Foundation.
References

[1] I. Mehdizadeh Gohari, M.A. Navarro, J. Li, A. Shrestha, F. Uzal, B.A. McClane, Pathogenicity and virulence of *Clostridium perfringens*. Virulence 12 (2021) 723–753.

[2] E. Scallan, R.M. Hoekstra, F.J. Angulo, R.V. Tauxe, M.A. Widdowson, S.L. Roy, J.L. Jones, P.M. Griffin, Foodborne illness acquired in the United States—Major pathogens, Emerg. Infect. Dis. 17 (2011) 7–15.

[3] E.C. Todd, Epidemiology of foodborne diseases: a worldwide review, World Health Stat. Q 50 (1997) 30–50.

[4] M. Richter, R. Rosselló-Móra, Shifting the genomic gold standard for the prokaryotic species definition, Proc. Natl. Acad. Sci. USA 106 (2009) 19126–19131.

[5] F.A. Uzal, J.C. Freedman, A. Shrestha, J.R. Theoret, J. Garcia, M.M. Awad, V. Adams, R.J. Moore, J.I. Rood, B.A. McClane, Towards an understanding of the role of *Clostridium perfringens* toxins in human and animal disease, Future Microbiol. 9 (2014) 361–377.

[6] L.M. Marler, J.A. Siders, L.C. Wolters, Y. Pettigrew, B.L. Skitt, S.D. Allen, Comparison of five cultural procedures for isolation of *Clostridium difficile* from stools, J. Clin. Microbiol. 30 (1992) 514–516.

[7] M. Baym, S. Kryazhimskiy, T.D. Liebermann, H. Chung, M.M. Desai, R. Kishony, Inexpensive multiplexed library preparation for megabase-sized genomes, PLoS One 10 (2015) e0128036.

[8] A. Bankevich, S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, S.I. Nikolenko, S. Pham, A.D. Prjibelski, A.V. Pyshkin, A.V. Sirotkin, N. Vyahhi, G. Tesler, M.A. Alekseyev, P.A. Pevzner, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, J. Comput. Biol. 19 (2012) 455–477.

[9] T. Brettin, J.J. Davis, T. Disz, R.A. Edwards, S. Gerdes, G.J. Olsen, R. Overbeek, B. Parrello, G.D. Pusch, M. Shukla, J.A. Thomason 3rd, R. Stevens, V. Vonstein, A.R. Wattam, F. Xia, RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes, Sci. Rep. 5 (2015) 8365.

[10] K.T. Konstantinidis, J.M. Tiedje, Genomic insights that advance the species definition for prokaryotes, Proc. Natl. Acad. Sci. USA 102 (2005) 2567–2572.

[11] J.J. Davis, A.R. Wattam, R.K. Aziz, T. Brettin, R. Butler, R.M. Butler, P. Chlenski, N. Conrad, A. Dickerman, E.M. Dietrich, J.L. Gabbard, S. Gerdes, A. Guard, R.W. Kenyon, D. Machi, C. Mao, D. Murphy-Olson, M. Nguyen, E.K. Nordberg, G.J. Olsen, R.D. Olson, J.C. Overbeek, R. Overbeek, B. Parrello, G.D. Pusch, M. Shukla, C. Thomas, M. VanOeffelen, V. Vonstein, A.S. Warren, F. Xia, D. Xie, H. Yoo, R. Stevens, The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities, Nucleic Acids Res. 48 (2020) D606–D612.

[12] A. Carattoli, E. Zankari, A. García-Fernández, M. Vollby Larsen, O. Lund, L. Villa, F. Moller Aarestrup, H. Hasman, In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing, Antimicrob. Agents Chemother. 58 (2014) 3895–3903.

[13] Y. Feng, S. Zou, H. Chen, Y. Yu, Z. Ruan, BacWGSdb 2.0: a one-stop repository for bacterial whole-genome sequence typing and source tracking, Nucleic Acids Res. 49 (2021) D644–D650.