Draft Genome Sequence of pyschrotolerant *Clostridium* sp. M14 Isolated from Spoiled Uncooked Venison.

Nikola Palevich[a], Faith P. Palevich[a], Paul H. Maclean[a], Ruy Jauregui[a], Eric Altermann[a,b], John Mills[a] and Gale Brightwell[a].

[a]AgResearch Limited, Grasslands Research Centre, Palmerston North, New Zealand.

[b]Riddet Institute, Massey University, Palmerston North, New Zealand.

Corresponding author: E-mail: nik.palevich@agresearch.co.nz.

ABSTRACT

The non-proteolytic, non-toxigenic and pyschrotolerant *Clostridium* sp. M14 was isolated from vacuum-packaged refrigerated spoiled venison. This report describes the generation and annotation of the 3.9 Mb draft genome sequence of *Clostridium* sp. M14.

ANNOUNCEMENT

The pyschrotolerant *Clostridium* sp. M14 is a Gram-positive, rod-shaped, spore-forming and slow-growing obligate anaerobe. M14 is able to grow at temperatures below 4°C, is catalase□ and oxidase□negative and metronidazole□sensitive, and originally isolated from vacuum-packaged refrigerated spoiled venison (1). Numerous bacterial species belonging to the *Clostridium* genera have been recognized as causative agents of blown pack spoilage (BPS) in vacuum packed meat products (2), and M14 was selected for genome sequencing to examine its role in BPS.

Previous studies identified a number of isolates with some similarity to that of non-proteolytic *C. botulinum* type B using classical bio-chemical differentiation, 16*S* rRNA gene Restriction

Fragment Length Polymorphism (RFLP) pattern analyses, 16S rRNA gene sequencing and PCR amplification of the ITS regions (1, 3). However, none of these isolates were deemed toxigenic. Broda *et al*., concluded that although the growth of such microorganisms in vacuum□packed chilled meat leads to product spoilage, it does not prejudice product safety. Here, we present a draft genome sequence of a representative meat spoilage associated non-proteolytic and non-toxigenic pyschrotolerant *Clostridium* sp. M14, that falls within meat associated pyschrotolerant *Clostridium* ARDRA Group 7 (4).

Strain M14 was isolated from a fully blown pack of vacuum packaged venison nearly 20 years ago and cultured anaerobically at 10°C in pre-reduced Peptone, Yeast Extract, Glucose, Starch broth (PYGS) as previously described (5). Genomic DNA was extracted using a modified phenol-chloroform procedure (5) and mechanically sheared using a Nebulizer instrument (Invitrogen) to select fragments of approximately 550 bp. A DNA library was prepared using the Illumina TruSeq™ Nano method and sequenced on the Illumina MiSeq platform with the 2× 250 bp paired-end (PE) reagent kit v2 producing a total of 3,125,724 PE raw reads. The quality of the raw reads was checked in FastQC v0.11.5 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), the reads were trimmed with Trimmomatic v0.39 (http://www.usadellab.org/cms/?page=trimmomatic) and assembled using the A5-miseq pipeline v20169825 with standard parameters (6). The *de novo* assembly of M14 produced 36 scaffolds with 184× coverage and an N50 value of 757,921 bp, with the largest scaffold length being 1,669,648 bp in size. The draft genome sequence is composed of 3,986,879 bp, with a %G+C content of 27.1%. A total of 3,717 putative protein-coding genes (PCGs) were predicted along with 81 tRNA, 19 rRNA, 170 ncRNA elements using GAMOLA2 (7). In addition, Diamond v0.9.21.122 (8) and InterProScan v5.36-75.0 (9) were used to search the

NCBI "nr" database with the resulting protein set imported into BLAST2GO as implemented in the OmicsBox software package v1.1.164 (10) where gene ontology terms and final annotations were assigned to each protein. All bioinformatics analyses were performed using default settings and parameters.

Carbohydrate-Active enZYme (11) profiling was analyzed using dbCAN2 (12) and revealed that the M14 genome is predicted to encode 44 glycoside hydrolases (GHs), 27 glycosyl transferases (GTs), 10 carbohydrate esterases (CEs) and 19 carbohydrate-binding protein module (CBM) families. Overall, approximately 2.7% of the M14 genome (100 CDSs) is predicted to encode either secreted or intracellular proteins dedicated to carbohydrate and even polysaccharide degradation. A comparison of the M14 with the recently characterized *C. estertheticum* subsp. *laramiense* type strain DSM 14864^{T} (ATCC 51254) (5), revealed clear differences in their enzymatic profiles. While both appear to be equipped to utilize many oligo- and monosaccharides as substrates for growth and encode a large repertoire of enzymes predicted to metabolize complex insoluble polysaccharides, M14 lacked genes encoding polysaccharide lyases (PLs).

The non-toxigenic status of M14 was confirmed by performing a bioinformatics search of the whole genome sequence (WGS) database (visited March 2020) at the National Center for Biotechnology Information (NCBI), using the predicted protein translation products of the known deadly neurotoxins of *C. botulinum* (BoNTs) protein sequences. None of the putative neurotoxin genes within the Group II *C. botulinum* BoNT operon that consists of *ha* (haemagglutinin) types (*ha70*, *ha17* and *ha34*) and accessory proteins *ntnh* (nontoxic-nonhaemagglutinin) and *botR* (neurotoxin regulator protein) upstream of the boNT gene, were identified in M14. In addition, a whole-genome alignment to the publicly available Group II *C.*

botulinum genomes revealed a low sequence homology with approximately <20% similarity to M14.

The genome sequence of the pyschrotolerant *Clostridium* sp. M14 reported here is a valuable resource for future studies investigating the bacterial genetic mechanisms associated with BPS. In order to improve the phylogenetic resolution of the *Clostridium* genera and improve our limited knowledge of meat spoilage caused by non-toxigenic *Clostridia* species.

Data availability. The genome sequence and associated data for *Clostridium* sp. M14 were deposited under GenBank accession number JAAMNF000000000, BioProject accession number PRJNA574489 and in the Sequence Read Archive (SRA) under accession number SRR11113219.

REFERENCES

1. Broda DM, Boerema JA, Bell RG. 1998. A PCR survey of psychrotrophic *Clostridium botulinum* like isolates for the presence of BoNT genes. *Letters in Applied Microbiology*, 27:219-223. **https://doi.org/10.1046/j.1472-765X.1998.00420.x**.
2. Mills J, Donnison A, Brightwell G. 2014. Factors affecting microbial spoilage and shelf-life of chilled vacuum-packed lamb transported to distant markets: A review. *Meat Science*, 98:71-80. **https://doi.org/10.1016/j.meatsci.2014.05.002**.
3. Broda DM, Boerema JA, Bell RG. 2003. PCR detection of psychrophilic Clostridium spp. causing 'blown pack' spoilage of vacuum-packed chilled meats. *Journal of Applied Microbiology*, 94:515-522. **https://doi.org/10.1046/j.1365-2672.2003.01859.x**.
4. Brightwell G, Horváth KM. 2018. Molecular discrimination of New Zealand sourced meat spoilage associated psychrotolerant Clostridium species by ARDRA and its comparison with 16s RNA gene sequencing. *Meat Science*, 138:23-27. **https://doi.org/10.1016/j.meatsci.2017.12.007**.
5. Palevich N, Palevich FP, Maclean PH, Jauregui R, Altermann E, Mills J, Brightwell G. 2019. Draft Genome Sequence of *Clostridium estertheticum* subsp. *laramiense* DSM 14864^T, Isolated from Spoiled Uncooked Beef. Microbiology Resource Announcements, 8:e01275-19. **https://doi.org/10.1128/MRA.01275-19**.
6. Coil D, Jospin G, Darling AE. 2014. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics* 31:587-589. **https://doi.org/10.1093/bioinformatics/btu661**.
7. Altermann E, Lu J, McCulloch A. 2017. GAMOLA2, a comprehensive software package for the annotation and curation of draft and complete microbial genomes. *Front Microbiol* 8:346. **https://doi.org/10.3389/fmicb.2017.00346**.
8. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59. **https://doi.org/10.1038/nmeth.3176**.
9. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30:1236-1240. **https://doi.org/10.1093/bioinformatics/btu031**.
10. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674-3676. **https://doi.org/10.1093/bioinformatics/bti610**.
11. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res* 37:D233-D238. **https://doi.org/10.1093/nar/gkn663**.
12. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. 2018. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 46:W95-W101. **https://doi.org/10.1093/nar/gky418**.