



Complete Genome Sequences of *Mycobacterium bovis* Strains Affiliated with Bovine Tuberculosis Outbreaks in Canada in 2016 and 2018

 Olga Andrievskaia,^a Amalia Garceac,^a Dara Lloyd,^a Mirjana Savic,^a  Marc-Olivier Duceppe^a

^aCanadian Food Inspection Agency, Ottawa Laboratory Fallowfield, Ottawa, Ontario, Canada

ABSTRACT *Mycobacterium bovis* is the primary causative agent of bovine tuberculosis, a zoonotic infectious disease that presents a risk to public health, livestock, and wildlife. Here, we report complete genome sequences of two *Mycobacterium bovis* strains affiliated with bovine tuberculosis outbreaks in Canadian cattle farms in 2016 and 2018.

Bovine tuberculosis (bTb) is an infectious zoonotic disease caused by *Mycobacterium bovis* and can present a risk to public health, the livestock industry, and wild animal species (1). Bovine tuberculosis is largely considered to be eliminated from Canadian livestock (2), with only two sporadic localized outbreaks detected in cattle farms over the past decade. Whole-genome sequencing (WGS) was performed for *M. bovis* strains affiliated with these outbreaks to provide references for epidemiological investigations and comparisons with *M. bovis* strains observed historically in Canada (3). The *M. bovis* strains 2016/0386 and 2018/0565 were isolated from mycobacteriosis-compatible granulomas collected from cattle in the provinces of Alberta in 2016 and British Columbia in 2018, respectively. *M. bovis* culture isolation and identification were performed at the National Reference Laboratory for Bovine Tuberculosis, Canadian Food Inspection Agency, according to standard microbiological methods (4). *M. bovis* isolates were stored in Proskauer-Beck medium supplemented with 5% horse serum at -80°C . For WGS, *M. bovis* isolates were grown on Middlebrook 7H11 agar for 4 weeks. Genomic DNA was purified using a combination of mechanical cell lysis with 0.5-mm ceramic beads (OMNI International, USA) and the MasterPure Gram-positive DNA purification kit (Lucigen, USA). Illumina sequencing libraries were prepared using Nextera XT DNA Library Preparation kit (Illumina Inc., USA) and sequenced on a MiSeq sequencer (Illumina Inc.) using a MiSeq reagent kit v2 to generate 2×250 -bp paired-end sequences. For Nanopore sequencing, DNA samples were treated with a Short Read Eliminator XS kit (Circulomics, USA). Nanopore libraries were prepared using the 1D Native barcoding genomic DNA protocol with EXP-NBD104 and SQK-LSK109 kits (Oxford Nanopore Technologies, UK) without shearing and sequenced using FLO-MIN111 flow cells on a MinION Mk1B device (Oxford Nanopore Technologies). The Nanopore basecalling was performed using Super Accuracy mode in Guppy v5.0.11, trimmed using Porechop v0.2.3, and filtered using Filtlong v0.2.1 (5). Long read assembly was performed with Tricycler (6) using Flye v2.7 (7), necat v0.0.1 (8), Canu v2.2 (9), and Shasta v0.8.0 (10). Reads were split in six subsamples. Circularized consensus sequences were corrected using Medaka v1.4.4 and polished with Illumina MiSeq reads using a combination of NextPolish v1.4.0 (11), ntEdit v1.3.5 (12), and Polypolish v0.5.0 (13) after trimming/filtering with fastp v0.23.2 (14). The sequencing coverage depth was determined using Minimap2 v2.17 (15) and SAMtools v1.7 (16) for long reads and BWA v0.7.17 and SAMtools for short reads. Default parameters of the above bioinformatics software were used unless otherwise specified (https://github.com/OLF-Bioinformatics/nanopore/blob/master/run_tricycler.sh;

Editor Steven R. Gill, University of Rochester School of Medicine and Dentistry

© Crown copyright 2023. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Olga Andrievskaia, Olga.Andrievskaia@inspection.gc.ca, or Marc-Olivier Duceppe, Marc-Olivier.Duceppe@inspection.gc.ca.

The authors declare no conflict of interest.

Received 15 December 2022

Accepted 31 January 2023

Published 14 February 2023

TABLE 1 Characteristics of complete genome sequences of two *Mycobacterium bovis* strains affiliated with bTb outbreaks in Canada in 2016 and 2018

M. bovis strain information		No. of reads by platform			Nanopore read N_{50} (bp)	Coverage (x)	Chromosome size (bp)	Predicted no. of CDSs ^a	Predicted no. of genes	Accession no. by database		
Name	Origin	Yr	Illumina	Nanopore						GenBank no.	SRA Nanopore	SRA Illumina
2016/0386	Alberta	2016	3,840,394	145,106	7,731	394	4,347,959	4,025	4,076	CPI12997	SRR22314236	SRR22314237
2018/0565	British Columbia	2018	14,581,052	61,840	5,842	576	4,350,529	4,022	4,073	CPI12996	SRR22314234	SRR22314235

^a CDSs, coding DNA sequences.

https://github.com/OLF-Bioinformatics/nanopore/blob/master/short_read_polish.sh). Gene prediction and annotation were done using NCBI Prokaryotic Genome Annotation Pipeline 2022-04-14.build6021 (17).

Assembly metrics and genomic characteristics are summarized in Table 1. The complete genomes of *M. bovis* strains 2016/0386 and 2018/0565 contain single chromosomes with a G+C content of 65.6%, 45 tRNAs, and 2 clustered regularly interspaced short palindromic repeats (CRISPRs). FastANI (v1.33) (18) whole-genome comparison of strains 2016/0386 and 2018/0565 with a reference *M. bovis* strain AF2122/97 (19) generated an average nucleotide identity of 99.98%.

Data availability. The assembled closed genome sequences of *M. bovis* strains 2016/0386 and 2018/0565 and raw reads are deposited in the NCBI GenBank under the BioProject number [PRJNA902023](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA902023). The accession numbers are listed in Table 1. The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

The study was funded by the Canadian Food Inspection Agency (CFIA GRDI project N-000195).

We acknowledge veterinary and technical staff involved in the collection of the diagnostic specimens.

REFERENCES

- Pérez-Lago L, Navarro Y, García-de-Viedma D. 2014. Current knowledge and pending challenges in zoonosis caused by *Mycobacterium bovis*: a review. *Res Vet Sci* 97:594–5100. <https://doi.org/10.1016/j.rvsc.2013.11.008>.
- Rossi G, Aubry P, Dubé C, Smith RL. 2019. The spread of bovine tuberculosis in Canadian shared pastures: data, model, and simulations. *Transbound Emerg Dis* 66:562–577. <https://doi.org/10.1111/tbed.13066>.
- Andrievskaia O, Duceppe MO, Lloyd D. 2018. Genome sequences of five *Mycobacterium bovis* strains isolated from farmed animals and wildlife in Canada. *Genome Announc* 6:e00258-18. <https://doi.org/10.1128/genomeA.00258-18>.
- Leao SC, Martin A, Mejia GI, Palomino JC, Robledo J, Telles M, Portaels F. 2004. Practical handbook for the phenotypic and genotypic identification of mycobacteria. Vanden Broele, Brugge, Belgium.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, Vezina B, Wyres KL, Holt KE. 2021. Tricycler: consensus long-read assemblies for bacterial genomes. *Genome Biol* 22:266. <https://doi.org/10.1186/s13059-021-02483-z>.
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>.
- Chen Y, Nie F, Xie SQ, Zheng YF, Dai Q, Bray T, Wang YX, Xing JF, Huang ZJ, Wang DP, He LJ, Luo F, Wang JX, Liu YZ, Xiao CL. 2021. Efficient assembly of nanopore reads via highly accurate and intact error correction. *Nat Commun* 12:60. <https://doi.org/10.1038/s41467-020-20236-7>.
- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
- Shafin K, Pesout T, Lorig-Roach R, Haukness M, Olsen HE, Bosworth C, Armstrong J, Tigyi K, Maurer N, Koren S, Sedlazeck FJ, Marschall T, Mayes S, Costa V, Zook JM, Liu KJ, Kilburn D, Sorensen M, Munson KM, Vollger MR, Monlong J, Garrison E, Eichler EE, Salama S, Haussler D, Green RE, Akeson M, Phillippy A, Miga KH, Carnevali P, Jain M, Paten B. 2020. Nanopore sequencing and the Shasta toolkit enable efficient de novo assembly of eleven human genomes. *Nat Biotechnol* 38:1044–1053. <https://doi.org/10.1038/s41587-020-0503-6>.
- Hu J, Fan J, Sun Z, Liu S. 2020. NextPolish: a fast and efficient genome polishing tool for long-read assembly. *Bioinformatics* 36:2253–2255. <https://doi.org/10.1093/bioinformatics/btz891>.
- Warren RL, Coombe L, Mohamadi H, Zhang J, Jaquish B, Isabel N, Jones SJ, Bousquet J, Bohlmann J, Birol IJB. 2019. ntEdit: scalable genome sequence polishing. *Bioinformatics* 35:4430–4432. <https://doi.org/10.1093/bioinformatics/btz400>.
- Wick RR, Holt KE. 2022. Polypolish: short-read polishing of long-read bacterial genome assemblies. *PLoS Comput Biol* 18:e1009802. <https://doi.org/10.1371/journal.pcbi.1009802>.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Li W, O'Neill KR, Haft DH, DiCuccio M, Chetverin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the Prokaryotic Genome Annotation Pipeline reach with protein family model curation. *Nucleic Acids Res* 49:1020–1028. <https://doi.org/10.1093/nar/gkaa1105>.
- Jain C, Rodriguez RL, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>.
- Malone KM, Farrell D, Stuber TP, Schubert OT, Aebersold R, Robbe-Austerman S, Gordon SV. 2017. Updated reference genome sequence and annotation of *Mycobacterium bovis* AF2122/97. *Genome Announcements* 5:e00157-17. <https://doi.org/10.1128/genomeA.00157-17>.