



Complete Genome Sequence of the Telford Type S Strain of *Mycobacterium avium* subsp. *paratuberculosis*

Rudiger Brauning,^a Karren Plain,^b Milan Gautam,^c Tonia Russell,^d C. Carolina Correa,^d Patrick Biggs,^e Richard Whittington,^b Alan Murray,^f Marian Price-Carter^g

^aAgResearch Ltd., Invermay Agricultural Centre, Mosgiel, New Zealand

^bUniversity of Sydney, Camden, New South Wales, Australia

^cSchool of Veterinary Science, EpiCentre, Massey University, Palmerston North, New Zealand

^dRamaciotti Centre for Genomics, University of New South Wales (UNSW), Sydney, Australia

^eInfectious Disease Research Centre, School of Veterinary Science, Hopkirk Research Institute, and School of Fundamental Sciences, Massey University, Palmerston North, New Zealand

^fSchool of Veterinary Science, Hopkirk Research Institute, Massey University, Palmerston North, New Zealand

^gAgResearch Ltd., Hopkirk Research Institute, Palmerston North, New Zealand

ABSTRACT *Mycobacterium avium* subsp. *paratuberculosis* is the causative agent of Johne's disease (JD). Here, we report the complete genome sequence of Telford 9.2, a well-characterized representative strain of the *M. avium* subsp. *paratuberculosis* S subtype that is endemic in New Zealand and Australian sheep.

Mycobacterium avium subsp. *paratuberculosis* is the causative agent of Johne's disease (JD), a chronic, generally subclinical but sometimes fatal granulomatous enteritis of ruminants (1). *M. avium* subsp. *paratuberculosis* subtype S (also called either subtype I or subtype III) has been isolated primarily from sheep but also from other ruminant species (reviewed in reference 2). Only draft genomes (3–5) of *M. avium* subsp. *paratuberculosis* type S are currently available. Here, we announce the complete genome sequence of Telford 9.2, an IS1311 type S IS900 restriction fragment length polymorphism (RFLP) type S1 strain. This is a clonal culture (passage level 5, including its primary isolation from sheep feces) of an isolate from a clinically infected sheep from New South Wales, Australia. It has been used as inoculum in an experimental model for clinical JD in sheep (6, 7), characterized genetically (8), and is representative of the *M. avium* subsp. *paratuberculosis* type endemic in Australian and New Zealand (NZ) sheep (9–11).

For Illumina and PacBio sequencing, bacterial stock was inoculated into either supplemented Middlebrook 7H9 (12) (Illumina) or M7H9C (PacBio) (13) liquid medium, cultured for 3 to 4 weeks, and then cultivated on modified Middlebrook 7H10 solid medium (12), harvested, and stored at -80°C .

Genomic DNA was prepared for both PacBio and Illumina sequencing by isopropanol precipitation and 70% ethanol wash of cetyltrimethylammonium bromide (CTAB)/phenol-chloroform-extracted cellular material after stepwise enzymatic digestion with lysozyme, RNase A, and proteinase K. For PacBio sequencing, the DNA was also digested with mutanolysin prior to proteinase K digestion and subjected to extra cleanup and concentration on Ampure PB magnetic beads.

A PacBio library was constructed and sequenced at the Ramaciotti Centre in Sydney, Australia, using P6-C4 chemistry, and sequenced to a coverage depth of $80\times$ on the PacBio RS II platform on a single-molecule real-time (SMRT) cell. It was improved with Illumina MiSeq 250-bp paired-end (PE) reads generated by sequencing two cultures of the Telford 9.2 reference strain. MiSeq-indexed libraries were created at New Zealand Genomics Limited using Nextera XT DNA kits (Illumina, San Diego, CA). Average

Citation Brauning R, Plain K, Gautam M, Russell T, Correa CC, Biggs P, Whittington R, Murray A, Price-Carter M. 2019. Complete genome sequence of the Telford type S strain of *Mycobacterium avium* subsp. *paratuberculosis*. Microbiol Resour Announc 8:e00004-19. <https://doi.org/10.1128/MRA.00004-19>.

Editor Christina Cuomo, Broad Institute of MIT and Harvard University

Copyright © 2019 Brauning et al. 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 Marian Price-Carter, Marian.Price-Carter@agresearch.co.nz.

Received 3 January 2019

Accepted 8 February 2019

Published 14 March 2019

TABLE 1 Discrepancies between Illumina and PacBio data^a

Position before fix	Variant type	Accepted solution	PacBio allele	Illumina allele	Applied fix
780880 ^b	Indel	Illumina	T	TG	Insertion
931746	SNP	PacBio	C	G	na
1112469	Indel	Illumina	GCCCCC	GCCCCC	Insertion
1302183	Indel	Illumina	AGGGG	AGGGGG	Insertion
1969375	Indel	Illumina	GCCCCC	GCCCCC	Insertion
2128150	Indel	Illumina	ACCCCC	ACCCCC	Insertion
2276090	Indel	Illumina	CGGGGG	CGGGGGG	Insertion
2577759	SNP	PacBio	G	A	na
2635929	Indel	Illumina	GCCCC	GCCCCC	Insertion
2642118	SNP	PacBio	C	T	na
2705636	Indel	Illumina	GCCCCC	GCCCCC	Insertion
3024648	Indel	Illumina	T	TC	Insertion
3201490	SNP	PacBio	C	G	na
3201602	SNP	PacBio	A	G	na
3211597	Indel	Illumina	CGGGGGGG	CGGGGGGGG	Insertion
3450836	Indel	PacBio	CATCGTCGCGCCGTGCTGGGCGCCAGCGCTCGC CGACCAGGCTGCGCGCCGCTCGACGCG CCGCGCGGCCCGCAGCGCTGCTGGG	C	na
4313098	SNP	Illumina	N	G	Base change
4314018	Indel	Illumina	GTTT	GTT	Deletion
4318473	Indel	Illumina	AC	A	Deletion
4319018	Indel	Illumina	AC	A	Deletion
4319236	Indel	Illumina	GTTT	GTT	Deletion
4319286	Indel	Illumina	CGGGG	CGGG	Deletion
4320148	Indel	Illumina	ACGCGCGC	ACGCGC	Deletion
4371898	SNP	PacBio	G	T	na
4416918	Indel	PacBio	CCGTTCCGGCGCCGAGCGTCACGCCAGCGTGCCGCTCG CGGGCCGGCGCCACGCTGGCGTGACG	CCG	na
4421523	Indel	Illumina	GCCCC	GCCCCC	Insertion
4572001	SNP	PacBio	G	A	na
4594338	Indel	Illumina	ACCCC	ACCCCC	Insertion

^a Illumina reads were mapped onto the PacBio assembly using BWA-MEM (17) v0.7.17-r1188 with parameter “-M,” and then variants (SNPs and indels) were detected (SAMtools [18] v1.3 with parameters “view -q 30 -F 256,” SAMtools v1.3 with parameters “mpileup -t DP,AD,” BCFtools v0.1.16 with parameters “call -cv,” BCFtools v0.1.16 with parameters “view -M2”). For each variant, a read depth greater than 10 was required, and a visual check of mapq values as well as the reference and alternative allele counts was performed. As a result of this analysis, for SNPs the PacBio alleles were accepted, for short indels the Illumina alleles were accepted, and for longer indels the PacBio alleles were accepted. All variants were verified by comparing 200 bp of flanking sequence (centered on the variants) to very closely related map strains (3, 4) using the “map to a reference” function in Geneious (19) and also comparing this fragment to *M. avium* subsp. *paratuberculosis* strains included in NCBI taxid 1770 using the NCBI BLAST service with default settings. SNP, single nucleotide polymorphism; na, no action.

^b For the indel at position 780880, the Telford1 sequence differed from closely related strains in both PacBio and Illumina alleles; Sanger sequencing confirmed the Illumina call.

coverage was 120× from PacBio data and 135× from Illumina data. There were 2.5 million Illumina PE reads (909 Mbp) and 150,000 PacBio reads prefilter (555 Mbp; N_{50} value, 10.5 kbp). PacBio reads went through default filtering steps in SMRTPipe v1.87.139483, which reduced read numbers to 63,000 (491 Mbp; N_{50} value, 10.8 kbp), and were assembled using PacBio Hierarchical Genome Assembly Process v3 (HGAP3; SMRT analysis v2.3.0) into a single contig (Telford1) of the size expected for a complete *M. avium* subsp. *paratuberculosis* genome and with a GC content of 69.2%, which is typical of *M. avium* subsp. *paratuberculosis* (3–5, 14). The PacBio-based assembly was improved by removing a 9-kbp overlap between the start and the end of the genome, orienting the genome with the start position at the beginning of the *dnaA* gene and mapping Illumina reads onto the PacBio assembly to detect and repair small-scale variations, as described in Table 1.

Telford1 has a sequence length of 4,907,428 bases, 4,377 coding sequences as predicted with the NCBI Prokaryotic Genome Annotation Pipeline (15), and an *in silico* IS1311 type S IS900 RFLP type S1 type (16).

Data availability. The genome assembly is available at GenBank under accession number CP033688 and the BioProject accession number PRJNA504678; raw data are available under SRA accession numbers SRX4997502 (Illumina) and SRX4997501 (PacBio), and *in silico* typing results can be found at <https://doi.org/10.6084/m9.figshare.7635977>.

ACKNOWLEDGMENTS

We thank Rebecca Maurer for her technical assistance with extraction of DNA for PacBio analysis.

Funding for this work was from a Meril SAS grant, Massey University project number RM16643.

REFERENCES

- Garcia AB, Shalloo L. 2015. Invited review: the economic impact and control of paratuberculosis in cattle. *J Dairy Sci* 98:5019–5039. <https://doi.org/10.3168/jds.2014-9241>.
- Stevenson K. 2015. Genetic diversity of *Mycobacterium avium* subspecies *paratuberculosis* and the influence of strain type on infection and pathogenesis: a review. *Vet Res* 46:64. <https://doi.org/10.1186/s13567-015-0203-2>.
- Mobius P, Holzer M, Felder M, Nordsiek G, Groth M, Kohler H, Reichwald K, Platzer M, Marz M. 2015. Comprehensive insights in the *Mycobacterium avium* subsp. *paratuberculosis* genome using new WGS data of sheep strain JIII-386 from Germany. *Genome Biol Evol* 7:2585–2601. <https://doi.org/10.1093/gbe/evv154>.
- Bannantine JP, Wu CW, Hsu C, Zhou S, Schwartz DC, Bayles DO, Paustian ML, Alt DP, Sreevatsan S, Kapur V, Talaat AM. 2012. Genome sequencing of ovine isolates of *Mycobacterium avium* subspecies *paratuberculosis* offers insights into host association. *BMC Genomics* 13:89. <https://doi.org/10.1186/1471-2164-13-89>.
- Wynne JW, Bull TJ, Seemann T, Bulach DM, Wagner J, Kirkwood CD, Michalski WP. 2011. Exploring the zoonotic potential of *Mycobacterium avium* subspecies *paratuberculosis* through comparative genomics. *PLoS One* 6:e22171. <https://doi.org/10.1371/journal.pone.0022171>.
- Begg DJ, de Silva K, Di Fiore L, Taylor DL, Bower K, Zhong L, Kawaji S, Emery D, Whittington RJ. 2010. Experimental infection model for Johne's disease using a lyophilised, pure culture, seedstock of *Mycobacterium avium* subspecies *paratuberculosis*. *Vet Microbiol* 141:301–311. <https://doi.org/10.1016/j.vetmic.2009.09.007>.
- Dukkupati VSR, Ridler AL, Thompson KG, Buddle BM, Hedgespeth BA, Price-Carter M, Begg DJ, Whittington RJ, Gicquel B, Murray A. 2016. Experimental infection of New Zealand Merino sheep with a suspension of *Mycobacterium avium* subspecies *paratuberculosis* (Map) strain Telford: kinetics of the immune response, histopathology and Map culture. *Vet Microbiol* 195:136–143. <https://doi.org/10.1016/j.vetmic.2016.09.018>.
- Marsh IB, Bannantine JP, Paustian ML, Tizard ML, Kapur V, Whittington RJ. 2006. Genomic comparison of *Mycobacterium avium* subsp. *paratuberculosis* sheep and cattle strains by microarray hybridization. *J Bacteriol* 188:2290–2293. <https://doi.org/10.1128/JB.188.6.2290-2293.2006>.
- Verdugo C, Pleydell E, Price-Carter M, Prattley D, Collins D, de Lisle G, Vogue H, Wilson P, Heuer C. 2014. Molecular epidemiology of *Mycobacterium avium* subsp. *paratuberculosis* isolated from sheep, cattle and deer on New Zealand pastoral farms. *Prev Vet Med* 117:436–446. <https://doi.org/10.1016/j.prevetmed.2014.09.009>.
- Gautam M, Ridler A, Wilson PR, Heuer C. 2018. Control of clinical paratuberculosis in New Zealand pastoral livestock. *N Z Vet J* 66:1–8. <https://doi.org/10.1080/00480169.2017.1379914>.
- Whittington RJ, Hope AF, Marshall DJ, Taragel CA, Marsh I. 2000. Molecular epidemiology of *Mycobacterium avium* subsp. *paratuberculosis*: IS900 restriction fragment length polymorphism and IS1311 polymorphism analyses of isolates from animals and a human in Australia. *J Clin Microbiol* 38:3240–3248.
- Whittington RJ. 2010. Cultivation of *Mycobacterium avium* subspecies *paratuberculosis*, p 244–266. *In* Behr M, Collins D (eds), *Paratuberculosis: organism, disease, control*. CABI, Wallingford, United Kingdom.
- Whittington RJ, Whittington A-M, Waldron A, Begg DJ, de Silva K, Purdie AC, Plain KM. 2013. Development and validation of a liquid medium (M7H9C) for routine culture of *Mycobacterium avium* subsp. *paratuberculosis* to replace modified Bactec 12B medium. *J Clin Microbiol* 51:3993–4000. <https://doi.org/10.1128/JCM.01373-13>.
- Li L, Bannantine JP, Zhang Q, Amonsin A, May BJ, Alt D, Banerji N, Kanjilal S, Kapur V. 2005. The complete genome sequence of *Mycobacterium avium* subspecies *paratuberculosis*. *Proc Natl Acad Sci U S A* 102:12344–12349. <https://doi.org/10.1073/pnas.0505662102>.
- NCBI. 2013. NCBI Prokaryotic Genome Annotation Pipeline. https://www.ncbi.nlm.nih.gov/genome/annotation_prok/.
- Price-Carter M, Whittington RJ. 2019. In silico typing of Telford. Figshare <https://doi.org/10.6084/m9.figshare.7635977.v1>.
- Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26:589–595. <https://doi.org/10.1093/bioinformatics/btp698>.
- Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27:2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>.