

Draft Genome Sequence of *Mycobacterium bovis* Strain SP38, a Pathogenic Bacterium Isolated from a Bovine in Brazil

Ana Marcia Sa Guimaraes,^a Cristina Kraemer Zimpel,^a Cássia Yumi Ikuta,^a Naíla Cannes do Nascimento,^b Andrea Pires dos Santos,^b Joanne Belle Messick,^b Marcos Bryan Heinemann,^a José Soares Ferreira Neto,^a Paulo Eduardo Brandão^a

Department of Preventive Veterinary Medicine and Animal Health, University of São Paulo, São Paulo, Brazil^a; Department of Comparative Pathobiology, Purdue University, West Lafayette, Indiana, USA^b

A.M.S.G. and C.K.Z. contributed equally to this work.

We report a draft genome sequence of *Mycobacterium bovis* strain SP38, isolated from the lungs of a cow in Brazil. The assembly of reads resulted in 36 contigs in a total of approximately 4.37 Mb. Comparison of *M. bovis* strains sequenced to date will aid in understanding bovine tuberculosis in Brazil.

Received 14 April 2015 Accepted 20 April 2015 Published 21 May 2015

Citation Guimaraes AMS, Zimpel CK, Ikuta CY, do Nascimento NC, dos Santos AP, Messick JB, Heinemann MB, Ferreira Neto JS, Brandão PE. 2015. Draft genome sequence of *Mycobacterium bovis* strain SP38, a pathogenic bacterium isolated from a bovine in Brazil. *Genome Announc* 3(3):e00511-15. doi:10.1128/genomeA.00511-15.

Copyright © 2015 Guimaraes et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Ana Marcia Sa Guimaraes, anamarcia@usp.br, or Cristina Kraemer Zimpel, criskzimpel@usp.br.

Mycobacterium bovis is the causative agent of bovine tuberculosis, an OIE (World Organisation for Animal Health) notifiable disease, responsible for major economic losses in livestock. Bovine tuberculosis is a contagious, infectious disease that can affect several domestic and wild animals besides human beings. In Brazil, the last evaluation of the National Program for the Control and Eradication of Animal Tuberculosis indicated a low prevalence of bovine tuberculosis (~1.3%). The prevalence of *M. bovis* infection in humans, however, is unknown. To better understand the bacterial maintenance in diverse populations and the transmission involving several animal species, it is necessary to comprehend the metabolism and mechanisms of pathogenicity of national *M. bovis* strains by comparing them to other sequenced strains. Therefore, the goal of the present study was to sequence and annotate the complete genome of *M. bovis* SP38, a Brazilian strain isolated from an infected bovine.

M. bovis was isolated from the lungs of an infected cow from a slaughterhouse in Sao Paulo, Brazil, following standard procedures (1) (IACUC protocol 1628/2009). DNA was extracted according to previous protocols (2, 3) and whole-genome sequencing from a paired-end library (TruSeq DNA sample preparation kit; Illumina, CA, USA) was performed using Illumina HiSeq2500. Reads were assembled using ABySS version 1.2.7 and Projector2. First-pass annotation was obtained using the NCBI pipeline. BLASTclust (4) was used to identify paralogous gene families (PGF) and predicted protein-coding sequences (CDSs) were compared to a set of proteins previously identified in the Brazilian and United Kingdom PPD (purified protein derivative of tuberculin from *M. bovis* strain AN5) (5) using BLAST (6).

A total of 29,651,856 reads were assembled into 36 contigs, resulting in a genome size of ~4.37 Mb with a GC content of 65.6%. First-pass annotation resulted in 3,986 CDSs, 3 rRNAs, and 45 tRNAs. Of these CDSs, 1,059 (26.57%) were annotated as hypothetical proteins and the remaining 2,927 (73.43%) had putative functions. A total of 308 PGFs, with 2 to 124 protein members each, were identified, comprising a total of 1,089 CDSs

(27.3%). These numbers are similar to those obtained from other *M. bovis* strains deposited in GenBank. Orthologous CDSs to all 116 PPDs were identified in the *M. bovis* strain SP38 and *M. tuberculosis* strain H37Rv. Low protein-sequence variability was detected in 8 of the SP38 orthologs and in 27 of the H37Rv orthologs (98% to 100% identity, 100% coverage), while moderate variability was observed in only 1 SP38 ortholog (95% identity, 100% coverage) and in 3 H37Rv orthologs (89 to 95% identity, 48 to 98% coverage). Nevertheless, 43 (37%) of these proteins are grouped into 37 PGFs in the SP38 strain, and differential gene expression has been already described for some of these families (7–10). In conclusion, differential gene expression within PGFs should be investigated as a possible source of host-bacterium relationship diversity and tuberculin skin test variation. Further analyses of this genome will provide information about *M. bovis* SP38 biology compared to other *M. bovis* strains.

Nucleotide sequence accession numbers. The *Mycobacterium bovis* SP38 sequence was deposited in GenBank under the accession number [JXOQ00000000](https://ncbi.nlm.nih.gov/nucl/JXOQ00000000). The version described in this paper is version [JXOQ00000000.1](https://ncbi.nlm.nih.gov/nucl/JXOQ00000000).

ACKNOWLEDGMENTS

This work was supported by the Science without Borders Program, Ministry of Science, Brazil (grant ID A008_2013). A.M.S.G., C.K.Z., J.S.F.N., M.B.H., and P.E.B. are indebted to CNPq and CAPES for the fellowships.

We are grateful to the Genomic Core Facilities from Purdue University (West Lafayette, IN, USA).

REFERENCES

1. Centro Panamericano de Zoonosis (CPZ). 1985. Tuberculosis—Bacteriología de la tuberculosis: el cultivo del *Mycobacterium tuberculosis*, Nota Técnica, no. 27. CPZ, Buenos Aires.
2. Bemer-Melchior P, Drugeon HB. 1999. Inactivation of *Mycobacterium tuberculosis* for DNA typing analysis. *J Clin Microbiol* 37:2350–2351.
3. Van Soolingen D, De Haas PE, Hermans PW, Van Embden JD. 1994. DNA fingerprinting of *Mycobacterium tuberculosis*. *Methods Enzymol* 235:196–205.

4. Biegert A, Mayer C, Remmert M, Söding J, Lupas AN. 2006. The MPI bioinformatics toolkit for protein sequence analysis. *Nucleic Acids Res* 34:W335–W339. <http://dx.doi.org/10.1093/nar/gkl217>.
5. Borsuk S, Newcombe J, Mendum TA, Dellagostin OA, McFadden J. 2009. Identification of proteins from tuberculin purified protein derivative (PPD) by LC-MS/MS. *Tuberculosis* 89:423–430. <http://dx.doi.org/10.1016/j.tube.2009.07.003>.
6. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [http://dx.doi.org/10.1016/S0022-2836\(05\)80360-2](http://dx.doi.org/10.1016/S0022-2836(05)80360-2).
7. Akif M, Khare G, Tyagi AK, Mande SC, Sardesai AA. 2008. Functional studies of multiple thioredoxins from *Mycobacterium tuberculosis*. *J Bacteriol* 190:7087–7095. <http://dx.doi.org/10.1128/JB.00159-08>.
8. Bukka A, Price CT, Kernodle DS, Graham JE. 2011. *Mycobacterium tuberculosis* RNA expression patterns in sputum Bacteria indicate secreted Esx factors contributing to growth are highly expressed in active disease. *Front Microbiol* 2:266. <http://dx.doi.org/10.3389/fmicb.2011.00266>.
9. Fisher MA, Plikaytis BB, Shinnick TM. 2002. Microarray analysis of the *Mycobacterium tuberculosis* transcriptional response to the acidic conditions found in phagosomes. *J Bacteriol* 184:4025–4032. <http://dx.doi.org/10.1128/JB.184.14.4025-4032.2002>.
10. Schnappinger D, Ehrt S, Voskuil MI, Liu Y, Mangan JA, Monahan IM, Dolganov G, Efron B, Butcher PD, Nathan C, Schoolnik GK. 2003. Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: insights into the phagosomal environment. *J Exp Med* 198: 693–704. <http://dx.doi.org/10.1084/jem.20030846>.