Complete Genome Sequence of *Corynebacterium pseudotuberculosis* Strain PA01, Isolated from Sheep in Pará, Brazil

Jorianne T. C. Alves,[a] Adonney A. O. Veras,[a] Ana Lídia Q. Cavalcante,[a] Pablo H. C. G. de Sá,[a] Larissa M. Dias,[a] Luis C. Guimarães,[a] Ezequiel Morais,[b] André G. M. Silva,[b] Vasco Azevedo,[c] Rommel T. J. Ramos,[a] Artur Silva,[a] Adriana R. Carneiro[a]*

Federal University of Pará, Center of Genomics and System Biology, Laboratory of Genomic and Bioinformatics, Belém, Pará, Brazil[a]; Federal University of Pará, Campus of Castanhal, Pará, Brazil[b]; Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil[c]

* Present address: Adriana R. Carneiro, Instituto de Ciências Biológicas, Universidade Federal do Pará, Guamá, Belém, Pará, Brasil.

***Corynebacterium pseudotuberculosis* is the etiological agent of caseous lymphadenitis disease. In this work, we present the first complete genome sequence of *Corynebacterium pseudotuberculosis* strain PA01, isolated in northern Brazil from an infected sheep. The genome length is 2,337,920 bp, and 2,003 coding sequences (CDS), 12 rRNAs, and 49 tRNAs were predicted.**

Address correspondence to Adriana R. Carneiro, adrianarc@ufpa.br.

Corynebacterium pseudotuberculosis is a facultative intracellular Gram-positive bacterium that belongs to the CMNR group (*Corynebacterium*, *Mycobacterium*, *Nocardia*, and *Rhodococcus*) (1), which causes caseous lymphadenitis (CLA), an infectious disease that affects small ruminants, mainly sheep and goats, and is characterized by the formation of abscesses in the superficial lymph nodes and subcutaneous tissues (1, 2).

This disease causes economic loss due to the progressive reduction in weight gain, depreciated wool and skin, reduced milk production, and eventually death caused by toxemia of the infected animals. It impacts sheep and goat farming around the globe, especially in the United States, Australia, South Africa, and Brazil (1, 3, 4). In Brazil, the agribusiness of goat and sheep has increased, especially in the northeast (http://www.agricultura.gov.br) and, consequently, there is a high prevalence of CLA disease in the states of Bahia (5), Pernambuco (6), and Rio Grande do Sul (7).

However, despite the bacteria epidemiology, there are no reports of its isolation in northern Brazil. *C. pseudotuberculosis* strain PA01 was isolated from the lymph nodes of a sheep in Pará, Brazil. Biochemical identification was performed using the API Coryne kit (bioMérieux, USA) and the strain was characterized as biovar *ovis*. Though it is nitrate reductase negative (8), molecular biology confirmation was obtained by a PCR multiplex with *rpoB*, *16S*, and *pld* genes (9). The genome was sequenced by the Ion Torrent PGM platform using a fragment library, which produced 1,894,790 reads. After sequencing, the reads were evaluated for quality using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and filtered and trimmed with average Phred quality scores equal to or greater than 20 by the FastX toolkit (http://hannonlab.cshl.edu/fastx_toolkit/), followed by assembly using Mira (10), which generated 22 contigs. The Lasergene 11 Core Suite with the SeqMan Proo tool was used to reduce the number of contigs to 5. The scaffold was obtained using Mauve (11) with *C. pseudotuberculosis* strain FRC41 (NC_014329) as the reference genome. Artemis software was uti-

lized to edit and fill gaps (12). Automatic annotation was performed using Rapid Annotation using Subsystem Technology (RAST) (13). The prediction of rRNAs and tRNAs were performed using RNAmmer (14) and tRNAscan-SE (15), respectively. The identification of protein domains and families was performed by InterProScan (16). All coding sequences (CDS) were manually curated using Artemis (12), BLASTp (http://blast.ncbi.nlm.nih.gov/), and the UniProt (http://www.uniprot.org) database. The identification and validation of the pseudogene was done using CLC Genomics Workbench (http://www.clcbio.com/).

The *C. pseudotuberculosis* strain PA01 genome consists of a circular chromosome of 2,337,920 bp, with 52.18% G+C content, 2,003 CDS, 12 rRNA operons, 49 tRNAs, and 17 pseudogenes predicted.

Nucleotide sequence accession number. This genome project has been deposited in GenBank under the accession number CP013327.

ACKNOWLEDGMENTS

FUNDING INFORMATION

REFERENCES

1. **Dorella FA, Pacheco LGC, Oliveira SC, Miyoshi A, Azevedo V.** 2006. *Corynebacterium pseudotuberculosis*: microbiology, biochemical proper-

ties, pathogenesis and molecular studies of virulence. Vet Res **37**:201–218. http://dx.doi.org/10.1051/vetres:2005056.

2. **Baird GJ, Fontaine MC.** 2007. *Corynebacterium pseudotuberculosis* and its role in ovine caseous lymphadenitis. J Comp Pathol **137**:179–210. http://dx.doi.org/10.1016/j.jcpa.2007.07.002.

3. **Dorella F, Gala-garcia A, Pinto AC, Sarrouh B, Antunes CA, Ribeiro D, Aburjaile FF, Fiaux KK, Guimarães LC, Seyffert N, El-aouar RA, Silva R, Hassan SS, Castro TLP.** 2009. Progression of "OMICS" methodologies for understanding the pathogenicity of *Corynebacterium pseudotuberculosis*: the Brazilian experience. Comput Struct Biotechnol J **6**:1–7. http://dx.doi.org/10.5936/csbj.201303013.

4. **Seyffert N, Guimarães AS, Pacheco LGC, Portela RW, Bastos BL, Dorella F, Heinemann MB, Lage P, Gouveia MG, Meyer R, Miyoshi A, Azevedo V.** 2010. High seroprevalence of caseous lymphadenitis in Brazilian goat herds revealed by *Corynebacterium pseudotuberculosis* secreted proteins-based ELISA. Res Vet Sci **88**:50–55. http://dx.doi.org/10.1016/j.rvsc.2009.07.002.

5. **Costa LFM.** 2002. *Corynebacterium pseudotuberculosis*, the etiological agent of the caseous lymphadenitis in goats. Rev Ci Méd Biol Salvador **1**:105–115.

6. **Abreu SRO, Mota RA, Rosinha GMS, Forner O, Pinheiro JW, Pereira RRB, Castro RS, Elisei C, Soares CO, Araújo FR, Madureira RC.** 2008. Comparação genotípica de isolados de *Corynebacterium pseudotuberculosis* de caprinos e ovinos do sertão de Pernambuco. Pesq Vet Bras **28**:481–487. http://dx.doi.org/10.1590/S0100-736X2008001000007.

7. **Camargo EV, Barboza CS, Krewer C, Vargas APC, Cecim M, Leal MLR.** 2010. Isolamento de *Corynebacterium pseudotuberculosis* no sêmen de um carneiro na região central do Rio Grande DO Sul. Arq Inst Biol, São Paulo **77**:139–142.

8. **Çetinkaya B, Karahan M, Atil E, Kalin R, De Baere T, Vaneechoutte M.** 2002. Identification of *Corynebacterium pseudotuberculosis* isolates from sheep and goats by PCR. Vet Microbiol **88**:75–83. http://dx.doi.org/10.1016/S0378-1135(02)00089-5.

9. **Pacheco LGC, Pena RR, Castro TLP, Dorella FA, Bahia RC, Carminati R, Frota MNL, Oliveira SC, Meyer R, Alves FSF, Miyoshi A, Azevedo V.** 2007. Multiplex PCR assay for identification of *Corynebacterium pseudotuberculosis* from pure cultures and for rapid detection of this pathogen in clinical samples. J Med Microbiol **56**:480–486. http://dx.doi.org/10.1099/jmm.0.46997-0.

10. **Chevreux B, Pfisterer T, Drescher B, Driesel AJ, Müller WE, Wetter T, Suhai S.** 2004. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Res **14**:1147–1159. http://dx.doi.org/10.1101/gr.1917404.

11. **Darling ACE, Mau B, Blattner FR, Perna NT.** 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res **14**:1394–1403. http://dx.doi.org/10.1101/gr.2289704.

12. **Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16**:944–945. http://dx.doi.org/10.1093/bioinformatics/16.10.944.

13. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.** 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics **9**:75. http://dx.doi.org/10.1186/1471-2164-9-75.

14. **Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW.** 2007. RNAmmer: consistent annotation of rRNA genes in genomic sequences. Nucleic Acids Res **35**:3100–3108.

15. **Lowe TM, Eddy SR.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res **25**:955–964. http://dx.doi.org/10.1093/nar/25.5.0955.

16. **Zdobnov EM, Apweiler R.** 2001. InterProScan—an integration platform for the signature-recognition methods in InterPro. Bioinformatics **17**:847–848. http://dx.doi.org/10.1093/bioinformatics/17.9.847.