Whole-Genome Sequence of *Corynebacterium pseudotuberculosis* 262 Biovar *equi* Isolated from Cow Milk

Carlos Leonardo de A. Araújo,[a] Larissa M. Dias,[a] Adonney A. O. Veras,[a] Jorianne T. C. Alves,[a] Ana Lídia Q. Cavalcante,[a] Christopher G. Dowson,[c] Vasco Azevedo,[b] Rommel T. J. Ramos,[a] Artur Silva,[a] Adriana R. Carneiro[a]*

Center of Genomics and System Biology, Laboratory of Genomic and Bioinformatics, Federal University of Pará, Belém, Pará, Brazil[a]; Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil[b]; School of Life Sciences, University of Warwick, Coventry, United Kingdom[c]

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

We report the complete genome sequence of *Corynebacterium pseudotuberculosis* 262, isolated from a bovine host. *C. pseudotuberculosis* is an etiological agent of diseases with medical and veterinary relevance. The genome contains 2,325,749 bp, 52.8% G+C content, 2,022 coding sequences (CDS), 50 pseudogenes, 48 tRNAs, and 12 rRNAs.

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

*C*orynebacterium pseudotuberculosis is a Gram-positive, facultative intracellular, pleomorphic, nonsporulating, noncapsulated, nonmotile bacterium that is the etiological agent of caseous lymphadenitis (CLA) in small ruminants and pyogranulomatus reactions, ulcerative lymphangitis, and mastitic, necrotic, and ulcerative dermatitis in cattle, all of which are diseases with medical and veterinary relevance. *C. pseudotuberculosis* affects several species, including sheep, goat, horse, cattle, llama, alpaca, buffalo, and human. This organism has various survival mechanisms and uses many strategies to adapt to its environment. After infection, the bacteria become encapsulated within walled-off lesions from which they evade immune system-mediated destruction, giving rise to a state of persistence (1–3). The molecular determinants of *C. pseudotuberculosis* virulence have been described and enable the search for potential targets for the development of new vaccine candidates by "omics" methodologies (4–7).

According to their capability for nitrate reduction, the strains of *C. pseudotuberculosis* are divided into two biovars. The organisms that perform the reduction of nitrate are classified into biovar *equi*, most of which have been isolated from horses and cattle. Bacteria that cannot perform the reduction of nitrate belong to biovar *ovis*, frequently isolated from sheep and goat (8). However, in cattle there are reports of infection by both biovars (9).

Here, we report the genome sequencing of *Corynebacterium pseudotuberculosis* 262, the first strain belonging to biovar *equi* isolated from a bovine host. This strain has been deposited in a collection in Belgium.

C. pseudotuberculosis strain 262 was isolated from cow milk, and the genome sequencing was performed with an Ion Torrent PGM platform chip 318, with a fragment library. A total of 388,943,492 bp were produced, with 166× genomic coverage. Subsequently, the tool FastQC 0.11.4 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to evaluate the raw data, and FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) was used to remove the reads with quality below Phred 20. The

genome assembly was performed by Mira 4.0.2 (http://mira-assembler.sourceforge.net), which produced 29 contigs with an N_{50} of 333,604 bp. The manual curation was performed through CLC Genomics Workbench 8 and Artemis 16.0.0 software (10). Automatic genome annotation was performed using Rapid Annotations using Subsystem Technology 2.0 (RAST) (11), and manual curation was performed with Artemis software and the nonredundant protein databases Uniprot (http://www.uniprot.org/) and the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/). tRNAs and rRNAs were predicted using the software tRNAScan-SE 1.21 (12) and RNAmmer 1.2 (13), respectively. The plasticity of pathogenicity islands (PAIs) was assessed with the Pathogenicity Island Prediction Software 1.1 (PIPS) (14), using *C. glutamicum* strain ATCC 21831 (CP007722.1) as the reference genome, which identified 10 pathogenicity islands.

The *C. pseudotuberculosis* strain 262 genome contains 2,325,749 bp, a G+C content of 52.8%, 2,022 coding sequences (CDS), 50 pseudogenes, 48 tRNAs, and 12 rRNAs.

Nucleotide sequence accession number. This whole-genome shotgun project has been deposited at GenBank under accession number CP012022.

FUNDING INFORMATION

REFERENCES

1. **Dorella FA, Pacheco LG, Oliveira SC, Miyoshi A, Azevedo V**. 2006. *Corynebacterium pseudotuberculosis* microbiology, biochemical properties, pathogenesis and molecular studies of virulence. Vet Res **37**:201–218.
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. **Yeruham I, Elad D, Friedman S, Perl S**. 2003. *Corynebacterium pseudotuberculosis* infection in Israeli dairy cattle. Epidemiol Infect **131:**947–955. http://dx.doi.org/10.1017/S095026880300894X.

4. **D'Afonseca V, Moraes PM, Dorella FA, Pacheco LG, Meyer R, Portela RW, Miyoshi A, Azevedo V**. 2008. A description of genes of *Corynebacterium pseudotuberculosis* useful in diagnostics and vaccine applications. Genet Mol Res **7:**252–260. http://dx.doi.org/10.4238/vol7-1gmr438.

5. **Dorella FA, 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 TL, Marques WS, Ramos R, Carneiro A, de Sá P, Miyoshi A, Azevedo V, Silva A**. 2013. 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.

6. **Seyffert N, Silva RF, Jardin J, Silva WM, de Paula Castro TL, Tartaglia NR, de Oliveira SKT, Portela RW, Silva A, Miyoshi A, Le Loir Y, Azevedo V**. 2014. Serological proteome analysis of *Corynebacterium pseudotuberculosis i*solated from different hosts reveals novel candidates for prophylactics to control caseous lymphadenitis. Vet Microbiol **174:** 255–260. http://dx.doi.org/10.1016/j.vetmic.2014.08.024.

7. **Radusky LG, Hassan S, Lanzarotti E, Tiwari S, Jamal S, Ali J, Ali A, Ferreira R, Barh D, Silva A, Turjanski AG, Azevedo VA**. 2015. An integrated structural proteomics approach along the druggable genome of *Corynebacterium pseudotuberculosis* species for putative druggable targets. BMC Genomics **16**(Suppl 5):S9. http://dx.doi.org/10.1186/1471-2164-16-S5-S9.

8. **Hassan SS, Guimarães LC, Pereira Ude P, Islam A, Ali A, Bakhtiar SM, Ribeiro D, Rodrigues Dos Santos A, Soares Sde C, de C, Dorella F, Pinto AC, Schneider MP, Barbosa MS, Almeida S, Abreu V, Aburjaile F, Carneiro AR, Cerdeira LT, Fiaux K, Barbosa E, Diniz C, Rocha FS, Ramos RT, Jain N, Tiwari S, Barh D, Miyoshi A, Müller B, Silva A, Azevedo V**. 2012. Complete genome sequence of *Corynebacterium pseudotuberculosis* biovar *ovis* strain p54b96 isolated from antelope in South Africa obtained by rapid next generation sequencing technology. Stand Genomic Sci **7:**189–199. http://dx.doi.org/10.4056/sigs.3066455.

9. **Yeruham I, Braverman Y, Shpigel NY, Chizov-Ginzburg A, Saran A, Winkler M**. 1996. Mastitis in dairy cattle caused by *Corynebacterium pseudotuberculosis* and the feasibility of transmission by houseflies I. Vet Q **18:**87–89. http://dx.doi.org/10.1080/01652176.1996.9694623.

10. **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.

11. **Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R**. 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res **42:** D206–D214. http://dx.doi.org/10.1093/nar/gkt1226.

12. **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.

13. **Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW**. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res **35:**3100–3108. http://dx.doi.org/10.1093/nar/gkm160.

14. **Soares SC, Abreu VAC, Ramos RTJ, Cerdeira L, Silva A, Baumbach J, Trost E, Tauch A, Hirata R, Jr, Mattos-Guaraldi AL, Miyoshi A, Azevedo V**. 2012. PIPS: pathogenicity island prediction software. PLoS One **7:**e30848. http://dx.doi.org/10.1371/journal.pone.0030848.