

# Metagenomic Assembly and Draft Genome Sequence of an Uncharacterized *Prevotella* sp. from Nelore Rumen

Luciano T. Kishi,<sup>a</sup> Raphael B. de Jesus,<sup>b</sup> Claudio D. Pavani,<sup>b</sup> Eliana G. M. Lemos,<sup>a</sup> Jackson A. M. de Souza<sup>b</sup>

UNESP, Univ Estadual Paulista, Departamento de Tecnologia/LBMP<sup>a</sup>; Departamento de Biologia/LGA, Jaboticabal, SP, Brazil<sup>b</sup>

***Prevotella* is one of the most abundant genera in bovine rumen, although no genome has yet been assembled by a metagenomics approach applied to Brazilian Nelore. We report the draft genome sequence of *Prevotella* sp., comprising 2,971,040 bp, obtained using the Illumina sequencing platform. This genome includes 127 contigs and presents a low 48% GC.**

Received 26 May 2015 Accepted 4 June 2015 Published 9 July 2015

**Citation** Kishi LT, de Jesus RB, Pavani CD, Lemos EGM, de Souza JAM. 2015. Metagenomic assembly and draft genome sequence of an uncharacterized *Prevotella* sp. from Nelore rumen. *Genome Announc* 3(4):e00723-15. doi:10.1128/genomeA.00723-15.

**Copyright** © 2015 Kishi 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 Jackson A. M. de Souza, [jackson@fcav.unesp.br](mailto:jackson@fcav.unesp.br).

*Prevotella* species have been isolated from a number of environments, including the human oral cavity (1), human skin abscesses (2), and soils (3). *Prevotella* species are also prevalent within the rumen and gastrointestinal tracts of herbivores and omnivores. The bacterial isolates used to establish the ruminal *Prevotella* group were originally described by Bryant et al. (4) on the basis of the nutritional requirement for hemin and were initially divided into two subspecies, *Bacteroides ruminicola* subsp. *ruminicola* (type strain 23) and *B. ruminicola* subsp. *brevis* (type strain GA33).

Various ruminal bacteria, including *Prevotella ruminicola* 23, have been shown to have esterase activity (5). However, knowledge regarding the diversity of esterases present in the genome, including the regulation, biochemical characteristics, and mechanistic action of these enzymes, is limited. Recently, the complete genome sequence of *P. ruminicola* 23 was made available (6), therefore providing an opportunity to evaluate the repertoire of enzymes enabling this organism to function as a highly efficient hemicellulose-degrading bacterium.

We collected ruminal content from fistulated 24-month-old Nelore cattle. Total DNA was extracted using the FastDNA SPIN kit (MP Biomedical, LLC). DNA sequencing was performed using the Illumina HiSeq2500 platform. *De novo* assembly was carried out using IDBA\_UD (7). Binning of the assembled contigs was carried out based on metagenomic read coverage, tetranucleotide frequency, and the occurrence of unique marker genes using MaxBin (8), which generated 127 clustered contigs. Our search of the 127 contigs against all bacterial and archaeal genome sequences available in the GenBank database (January 2015) using BLASTn (9) showed best hits for *P. ruminicola* 23. The sequence composition-based binning, the sequence homology search, and the metagenomic read recruitment collectively indicated that the 127 contigs represent well the draft genome sequence of the single-species population.

The draft genome is 2,971,040 bp, with 48% GC content ( $N_{50}$  value, 109,266), and contains a total of 2,515 protein-coding genes, 1 rRNA, and 48 tRNAs. Gene prediction and functional annotation of the draft genome were performed using the RAST

server (10). Further analysis of carbohydrate-active enzymes using the dbCAN database (11) have indicated the presence of 39 glycoside hydrolases, 43 glycosyl transferases, 4 polysaccharide lyases, 20 carbohydrate esterases, 6 auxiliary activities, and 25 carbohydrate-binding modules.

Functional genomics research is rapidly emerging, and its application via a metagenomics approach has revealed an important tool for the discovery and development of new and alternative pathways for biotechnology purposes. Our advances in understanding the biology and ecology of ruminal *Prevotellaceae* will significantly improve future applications for ruminant growth, nutrition, and even plant biomass conversion technologies.

**Nucleotide sequence accession number.** The sequences were deposited in the NCBI with accession number [SRX818104](https://www.ncbi.nlm.nih.gov/nuclseq/SRX818104).

## ACKNOWLEDGMENTS

We are grateful to T. T. Berchielli from the Department of Animal Science (UNESP) for bovine disposal.

PROPE-UNESP and the CAPES programs supported this work.

## REFERENCES

- Downes J, Tanner AC, Dewhirst FE, Wade WG. 2010. *Prevotella saccharolytica* sp. nov., isolated from the human oral cavity. *Int J Syst Evol Microbiol* 60:2458–2461. <http://dx.doi.org/10.1099/ijs.0.014720-0>.
- Sakamoto M, Ohkusu K, Masaki T, Kako H, Ezaki T, Benno Y. 2007. *Prevotella pleuritidis* sp. nov., isolated from pleural fluid. *Int J Syst Evol Microbiol* 57:1725–1728. <http://dx.doi.org/10.1099/ijs.0.64885-0>.
- Ueki A, Akasaka H, Satoh A, Suzuki D, Ueki K. 2007. *Prevotella paludivivens* sp. nov., a novel strictly anaerobic, Gram-negative, hemicellulose-decomposing bacterium isolated from plant residue and rice roots in irrigated rice-field soil. *Int J Syst Evol Microbiol* 57:1803–1809. <http://dx.doi.org/10.1099/ijs.0.64914-0>.
- Bryant MP, Small N, Bouma C, Chu H. 1958. *Bacteroides ruminicola* n. sp. and *Succinimonas amylolytica*; the new genus and species; species of succinic acid-producing anaerobic bacteria of the bovine rumen. *J Bacteriol* 76:15–23.
- Hespell RB, O'Bryan-Shah PJ. 1988. Esterase activities in *Butyrivibrio fibrisolvens* strains. *Appl Environ Microbiol* 54:1917–1922.
- Purushe J, Fouts DE, Morrison M, White BA, Mackie RI, North American Consortium for Rumen Bacteria, Coutinho PM, Henrissat B, Nelson KE. 2010. Comparative genome analysis of *Prevotella ruminicola*

- and *Prevotella bryantii*: insights into their environmental niche. *Microb Ecol* 60:721–729. <http://dx.doi.org/10.1007/s00248-010-9692-8>.
7. Peng Y, Leung HC, Yiu SM, Chin FY. 2012. IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28:1420–1428. <http://dx.doi.org/10.1093/bioinformatics/bts174>.
  8. Wu Y.-W, Tang Y.-H, Tringe SG, Simmons BA, Singer SW. 2014. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome* 2:26. <http://dx.doi.org/10.1186/2049-2618-2-26>.
  9. 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).
  10. 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. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.
  11. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. 2012. dbCAN: a Web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 40:W445–W451. <http://dx.doi.org/10.1093/nar/gks479>.