

Draft Genome Sequence of *Lactobacillus pasteurii* CRBIP 24.76^T

Sylvie Cousin,^a Dominique Clermont,^a Sophie Creno,^b Laurence Ma,^b Valentin Loux,^c Chantal Bizet,^a Christiane Bouchier^b

Institut Pasteur, Centre de Ressources Biologiques de l'Institut Pasteur, Département de Microbiologie, Paris, France^a; Institut Pasteur, Plate-forme Génomique, Département Génomes et Génétique, Paris, France^b; INRA, UR1077 Mathématique, Informatique et Génome, Jouy-en-Josas, France^c

We report the draft genome sequence of the type strain *Lactobacillus pasteurii* CRBIP 24.76, which is closely related to *L. gigeriorum* CRBIP 24.85^T, isolated from a chicken crop. The total length of the 29 contigs is about 1.9 Mb, with a G+C content of 40% and 1,946 coding sequences.

Received 26 July 2013 Accepted 29 July 2013 Published 22 August 2013

Citation Cousin S, Clermont D, Creno S, Ma L, Loux V, Bizet C, Bouchier C. 2013. Draft genome sequence of *Lactobacillus pasteurii* CRBIP 24.76^T. *Genome Announc.* 1(4):e00660-13. doi:10.1128/genomeA.00660-13.

Copyright © 2013 Cousin et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Christiane Bouchier, bouchier@pasteur.fr.

The genus *Lactobacillus* represents the largest group of facultatively anaerobic, catalase-negative, non-spore-forming, Gram-positive, rod-shaped organisms which produce lactic acid as a major end product of metabolism (1); the genus currently contains >150 recognized species or subspecies. Species of the genus *Lactobacillus* have been isolated from a wide range of habitats, such as the oral cavity, vagina, and gastrointestinal tracts of humans and animals, food, vegetation, and sewage, and are of great commercial importance due to their use in the production of a range of fermented dairy, meat, and vegetable products. *Lactobacillus* species are also the most widely used probiotics, aimed at promoting a healthy lifestyle. *Lactobacillus pasteurii* strain 1517^T was isolated in the early 1960s and deposited in the Collection de l'Institut Pasteur under the strain name CRBIP 24.76^T. Unfortunately, according to the depositor archives, the origin of the strain is not known. This strain represents a new species (2) that is well discriminated from its closest relative *Lactobacillus gigeriorum* CRBIP 24.85^T, isolated from a chicken crop (3).

Here, we report the draft genome sequence of *L. pasteurii* CRBIP 24.76^T, which was obtained using a whole-genome strategy based on Illumina paired-end sequencing with an insert length of 420 bp (Illumina genome analyzer HiSeq 2000). Quality-filtered reads (16,681,432 reads, 89 bases mean read length, ~730-fold coverage) were assembled using ABySS (version 1.2.6 [4]) with a k-mer of 40 bases. All reads were assembled into 30 contigs. The total length of the 30 contigs is 1,905,399 nucleotides (nt), with a G+C content of 40%. The length of the longest contig is 313,402 bases, and the smallest is 1,359 bases. The contigs were annotated with the AGMIAL platform (5), an integrated bacterial genome annotation system. The prediction of coding sequences used the self-training gene detection software SHOW based on hidden Markov models (<http://genome.jouy.inra.fr/ssb/SHOW/>). tRNAs and rRNAs were detected using tRNAscan-SE (6) and RNAmmer (7) softwares, respectively. The predicted conserved domain sites were 1,946 (1,917 complete), among which 11 rRNA operons with 3, 3, and 5 copies of 23S, 5S, and 16S, respectively, were detected, as well as 52 tRNA genes.

Nucleotide sequence accession numbers. The strain is publicly available in two European Biological Resource Centres under the no. CRBIP 24.76^T and DSM 23907^T. The draft of the whole-genome sequencing project has been deposited in EMBL under the accession no. [CAKD01000001](https://www.ebi.ac.uk/ena/browser/view/CAKD01000001) to [CAKD01000030](https://www.ebi.ac.uk/ena/browser/view/CAKD01000030). The version described in this paper is the first version.

ACKNOWLEDGMENTS

We thank N. Joly (Projects and Developments in Bioinformatics, Institut Pasteur, Paris) for the programs used for the quality filtering of the reads.

This communication is an initiative of the European Consortium of Microbial Resources Centres (EMbaRC), supported by the European Commission's Seventh Framework Programme (FP7, 2007–2013), Research Infrastructures action, under the grant agreement no. FP7-228310.

REFERENCES

1. Hammes WP, Hertel C. 2009. Genus I. *Lactobacillus*, p 465–511. In Vos P, Garrity G, Jones D, Krieg NR, Ludwig W, Rainey FA, Schleifer KH, Whitman WB (ed), *Bergey's manual of systematic bacteriology*, 2nd ed, vol 3. Springer-Verlag, New York, NY.
2. Cousin S, Motreff L, Gulat-Okalla ML, Gouyette C, Spröer C, Schumann P, Begaud E, Bouchier C, Clermont D, Bizet C. 2013. *Lactobacillus pasteurii* sp. nov. and *Lactobacillus hominis* sp. nov. *Int. J. Syst. Evol. Microbiol.* 63:53–59.
3. Cousin S, Gulat-Okalla ML, Motreff L, Gouyette C, Bouchier C, Clermont D, Bizet C. 2012. *Lactobacillus gigeriorum* sp. nov., isolated from chicken crop. *Int. J. Syst. Evol. Microbiol.* 62:330–334.
4. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I. 2009. ABySS: a parallel assembler for short read sequence data. *Genome Res.* 19:1117–1123.
5. Bryson K, Loux V, Bossy R, Nicolas P, Chaillou S, van de Guchte M, Penaud S, Maguin E, Hoebeke M, Bessières P, Gibrat JF. 2006. AGMIAL: implementing an annotation strategy for prokaryote genomes as a distributed system. *Nucleic Acids Res.* 34:3533–3545.
6. 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.
7. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108.