

Draft Genome Sequence of *Lactobacillus fermentum* Strain 3872

Andrey V. Karlyshev,^a Kavita Raju,^a Vyacheslav M. Abramov^b

School of Life Sciences, Faculty of Science, Engineering and Computing, Kingston University, Kingston upon Thames, United Kingdom^a; Institute of Immunological Engineering, Lyubuchany, Moscow Region, Russia^b

This report describes a draft genome sequence of *Lactobacillus fermentum* strain 3872. The data revealed remarkable similarity to and dissimilarity with the published genome sequences of other strains of the species. The absence of and variation in structures of some adhesins and the presence of an additional adhesin may reflect adaptation of the bacterium to different host systems and may contribute to specific properties of this strain as a new probiotic.

Received 25 October 2013 Accepted 27 October 2013 Published 27 November 2013

Citation Karlyshev AV, Raju K, Abramov VM. 2013. Draft genome sequence of *Lactobacillus fermentum* strain 3872. *Genome Announc.* 1(6):e01006-13. doi:10.1128/genomeA.01006-13.

Copyright © 2013 Karlyshev 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 Andrey V. Karlyshev, a.karlyshev@kingston.ac.uk.

Lactobacillus fermentum isolated from various sources, including humans, poultry, and pigs, was found to be beneficial as a probiotic (1). Application of *L. fermentum* affects body composition and metabolism, providing beneficial effects on human health (2). This bacterium has been used for treatment of gastrointestinal and respiratory illnesses (3). This study describes a draft genome sequence of *L. fermentum* strain 3872, isolated from breast milk from a healthy woman in Russia in 2011. This draft genome sequence is remarkably different from the complete genomes sequences of the other strains of this species (CECT 5716, IFO 3956, and F-6) available at the time of preparation of this article for publication. Despite the presence of highly conserved regions and almost identical rRNA genes, only 70% of reads could be mapped onto the reference genome sequences.

Sequencing using IonTorrent PGM and Chip 314 produced 414,245 reads, which were assembled using the IonTorrent assembler into 284 contigs over 1 kb in size, with a maximum length of 76,656 bases. The total size of the assembly (2.21 Mb, with 29.46-fold genome coverage) and GC content (50.9%) are in good agreement with figures for the reference strains (2.06 to 2.10 Mb and 51.5 to 51.7% for size and GC content ranges, respectively). Genome sequence annotation was conducted using the GenBank microbial genome annotation pipeline (NCBI) and RAST (4), which predicted 2,087 protein-encoding sequences (with the largest encoded protein being an ATP-dependent helicase, at 1,337 amino acids) and 2,060 protein-encoding sequences (with proteins of up to 1,447 amino acids long, for DNA polymerase III).

The test and reference genomes share a gene encoding colicin V (5), which may contribute to the probiotic properties of these bacteria. There are general multidrug resistance genes (e.g., encoding efflux pump proteins) and genes responsible for resistance to arsenate, heavy metals, tetracycline, beta-lactam antibiotics, and fluoroquinolones. There is a large number of transposable and insertion sequence (IS) elements (e.g., several copies of the IS200 transposase gene).

All strains also contain highly similar copies of a gene encoding a fibronectin-binding protein (up to 99% identity over the entire lengths of the proteins). Only a short C-terminal region of one

mucus-binding protein (MBP) found in all reference strains is present in the test strain, with another MBP completely missing from the latter. One outstanding feature of the genome of *L. fermentum* strain 3872 is the presence of a collagen-binding protein (CBP)-encoding gene completely missing from all three reference strains. This large (1,055-amino-acid-long) CBP has a number of long repeats in the C-terminal area and has the maximum level of similarity (up to 87% identity) to CBPs found in *L. paracasei*, *L. oris*, and *L. casei*. Compared to other CBPs, this protein of *L. fermentum* 3872 is unique due to the presence of an additional copy of a repeated element (ca. 80 amino acids in size), which may be responsible for binding to host cell receptors. The strain-to-strain variation in a repertoire of *L. fermentum* adhesins may reflect adaptation of this microorganism to specific host systems.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number [AVCT00000000](https://www.ncbi.nlm.nih.gov/nuccore/AVCT00000000). The version described in this paper is version AVCT01000000.

ACKNOWLEDGMENT

This work was not supported by any external funding.

REFERENCES

1. Lin WH, Yu B, Jang SH, Tsen HY. 2007. Different probiotic properties for *Lactobacillus fermentum* strains isolated from swine and poultry. *Anaerobe* 13:107–113.
2. Omar JM, Chan Y, Jones ML, Prakash S, Jones PJH. 2013. *Lactobacillus fermentum* and *Lactobacillus amylovorus* as probiotics alter body adiposity and gut microflora in healthy persons. *J. Funct. Foods* 5:116–123.
3. West NP, Pyne DB, Cripps AW, Hopkins WG, Eskesen DC, Jairath A, Christophersen CT, Conlon MA, Fricker PA. 2011. *Lactobacillus fermentum* (PCC) supplementation and gastrointestinal and respiratory-tract illness symptoms: a randomised control trial in athletes. *Nutr. J.* 10:30.
4. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formisano 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. doi:10.1186/1471-2164-9-75.
5. Abercrombie JG, Paynter MJB, Hayasaka SS. 2006. Ability of colicin V to control *Escherichia coli* O157:H7 in ground beef. *J. Food Saf.* 26:103–114.