

Draft Genome Sequence of *Lactobacillus crispatus* 2029

Andrey V. Karlyshev,^a Vyacheslav G. Melnikov,^b Valentin C. Khlebnikov,^c Vyacheslav M. Abramov^c

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

This report describes a draft genome sequence of *Lactobacillus crispatus* 2029. The reads generated by the Ion Torrent PGM were assembled into contigs with a total size of 2.2 Mb. The data were annotated using the NCBI GenBank and RAST servers. A comparison with the reference strain revealed specific features of the genome.

Received 15 December 2013 Accepted 29 January 2014 Published 20 February 2014

Citation Karlyshev AV, Melnikov VG, Khlebnikov VC, Abramov VM. 2014. Draft genome sequence of *Lactobacillus crispatus* 2029. *Genome Announc.* 2(1):e01221-13. doi:10.1128/genomeA.01221-13.

Copyright © 2014 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 crispatus strain 2029 was isolated in 2011 at the Institute of Engineering Immunology, Lyubuchany, Russia, upon investigation of a bacterial species spectrum of vaginal lactobacilli. Among 726 strains of *L. crispatus* tested, only 19 strains were able to produce hydrogen peroxide at a level of 100 to 120 mg/liter. Only three of these strains were found to be resistant to intestinal stress conditions. Of these, strain 2029 was the only one to satisfy the technological requirements. This strain was found to be able to adhere to vaginal and intestinal epithelial cells. Strain *L. crispatus* 2029 is deposited at the All-Russian Collection of Microorganisms at the G. K. Skryabin Institute of Biochemistry and Physiology of Microorganisms (Russian Academy of Sciences, Pushchino, Russia) under registration no. VKM B-2727D.

Here, we present a draft genome sequence of strain *L. crispatus* 2029. The only complete *L. crispatus* genome sequence available at the time of preparation of this paper was that of strain ST1 (1). The sequencing reads, generated by IonTorrent PGM using Chip 314, were assembled by Ion Torrent assembler into 295 contigs (2.0 to 40.8 kb). The total size of the assembly (2,193,123 bases, 36.08-fold genome coverage) and G+C content (36.9%) are in full agreement with the published data for other strains of *L. crispatus* (2.04 to 2.46 Mb and 36.7 to 37.1%).

The genome sequence was annotated using NCBI GenBank and RAST (2) genome annotation servers. Among the genes found were those required for bacteriocin biosynthesis and transport. Several S-layer protein-coding genes were detected, along with a gene encoding a fibronectin-binding protein (FBP). In contrast to *L. crispatus* ST1, a gene encoding a mucus binding protein was absent. Homologues to only four out of S-layer proteins, encoded by strain ST1, were present. A gene encoding a homologue of streptolysin S biosynthesis protein SagB was present in the test strain but not in the reference strain ST1.

One gene product was annotated by both programs as “lactocepin S-layer protein,” with a predicted size of 158 amino acids. A similar protein is also encoded by *L. crispatus* ST1 (158 amino acids [aa] long, 99% identity). In general, lactocepins are cell surface-located or -secreted peptidases responsible for anti-inflammatory effects via selective degradation of proinflammatory chemokines (3). However, these proteins are significantly larger in size, e.g., the originally de-

scribed lactocepin produced by *L. lactis* is 1,715 amino acids long (4). It remains to be elucidated whether the *L. crispatus* 2029 protein annotated as lactocepin has a similar function and activity.

The sequencing data were also analyzed by mapping reads onto the reference genome of strain ST1 using the CLC Genomics Workbench software. NCBI BLASTx analysis of assembled unmapped reads identified the genes present in the test strain genome but not in the reference genome. Among these genes were those encoding a pullulanase, phage-related proteins (integrase and resolvases), transposases, and a CRISPR-associated protein.

The draft genome sequence of *L. crispatus* described in this report will assist in the understanding of specific features of this strain as a promising probiotic.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [AVFH00000000](https://www.ncbi.nlm.nih.gov/nuccore/AVFH00000000). The version described in this paper is version AVFH02000000.

ACKNOWLEDGMENTS

This work was not supported by any external funding. The publication costs were covered by the International Science and Technology Center.

REFERENCES

- Ojala T, Kuparinen V, Koskinen JP, Alatalo E, Holm L, Auvinen P, Edelman S, Westerlund-Wikström B, Korhonen TK, Paulin L, Kankainen M. 2010. Genome sequence of *Lactobacillus crispatus* ST1. *J. Bacteriol.* 192:3547–3548. <http://dx.doi.org/10.1128/JB.00399-10>.
- 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, 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>.
- von Schille MA, Höermannsperger G, Weiher M, Alpert CA, Hahne H, Bäuerl C, van Huynegem K, Steidler L, Hrnčir T, Pérez-Martínez G, Kuster B, Haller D. 2012. Lactocepin secreted by *Lactobacillus* exerts anti-inflammatory effects by selectively degrading proinflammatory chemokines. *Cell Host Microbe* 11:387–396. <http://dx.doi.org/10.1016/j.chom.2012.02.006>.
- Christensson C, Pillidge CJ, Ward LJ, O'Toole PW. 2001. Nucleotide sequence and characterization of the cell envelope proteinase plasmid in *Lactococcus lactis* subsp. *cremoris* HP. *J. Appl. Microbiol.* 91:334–343. <http://dx.doi.org/10.1046/j.1365-2672.2001.01390.x>.