



Complete Genome Sequence of *Lactobacillus crispatus* CO3MRSI1

 Elias McComb,^a  Johanna Holm,^a Bing Ma,^a  Jacques Ravel^a

^aInstitute for Genome Sciences, University of Maryland Baltimore, Baltimore, Maryland, USA

ABSTRACT *Lactobacillus crispatus* is a commonly found bacterium in vertebrate microbiota, particularly the human vagina. We report the first complete genome of a strain isolated from a human vagina, *L. crispatus* CO3MRSI1.

Lactobacillus crispatus is frequently associated with vertebrate microbiota, particularly the human vaginal microenvironment (1). To date, the only closed genome to represent the species has been a strain isolated from a chicken crop (2). We report the first completed *Lactobacillus crispatus* genome sequence of a strain isolated from the human vagina. *L. crispatus* has often been associated with beneficial health outcomes such as decreased risk for sexually transmitted infection and lowered prevalence of pregnancy complications (3, 4). This genome will serve as a reference for studies aiming to better understand human-associated strains of *L. crispatus* and their roles in reproductive health.

L. crispatus CO3MRSI1 was isolated from a midvaginal swab collected from a participant in a study in Baltimore, MD, in 2010 (1). The University of Maryland Baltimore (UMB) institutional review board (IRB) approved the use of this sample. The swab was stored frozen in Amies transport medium, and then 50 μ l of thawed transport medium was plated onto a de Man-Rogosa-Sharpe (MRS) agar plate (Sigma-Aldrich, Carlsbad, CA) that was grown under anaerobic conditions in a Coy chamber (Grass Lakes, MI) at 37°C. After 48 h, a colony was picked and grown in MRS broth under the same conditions for 48 h. The DNA was extracted by removing the S-layer proteins (described by Johnson et al.) (5), followed by extraction using the MasterPure DNA purification kit (Lucigen, product no. MC85200) with two phenol-chloroform cleanups prior to DNA precipitation. The genomic DNA was sheared to 20 kb using needle shearing and was size selected using BluePippin (Sage Scientific, Beverly, MA). SMRTbell libraries were generated following standard library protocols of a DNA template preparation kit (Pacific Biosciences, Menlo Park, CA) and sequenced using one single-molecule real-time (SMRT) cell on a PacBio RS II system with the mode set at 7,750 bp. Library construction and sequencing were performed at the Institute for Genome Sciences at UMB.

Raw sequence reads (197,711 reads) were quality filtered and removed if they were less than 1,000 bp long and had less than 5% overlap. The average read length was 10,952 bp, resulting in 33.1 \times coverage of the genome. Reads were assembled using Canu v1.7 with default settings (6). The strain was identified as *L. crispatus* by means of phylogenetic analysis of both the full-length 16S rRNA gene sequence (99% sequence identity with the 16S rRNA genes in *L. crispatus* DSM 20584 and *L. crispatus* ATCC 33820) and marker genes using an adaptation of FastTree2 in Kbase (SpeciesTreeBuilder v0.0.12) (7). *L. crispatus* CO3MRSI1 consists of a single circular chromosome (2,345,902 bp) with a GC content of 37.1%. No plasmid was detected. Gene finding and annotation were performed on Kbase using Prokka v1.12 (8). The assembly contains 2,307 predicted coding sequences (CDS), 55% with putative function. The genome contains 5 rRNA operons and 64 tRNA

Citation McComb E, Holm J, Ma B, Ravel J. 2019. Complete genome sequence of *Lactobacillus crispatus* CO3MRSI1. Microbiol Resour Announc 8:e01538-18. <https://doi.org/10.1128/MRA.01538-18>.

Editor Frank J. Stewart, Georgia Institute of Technology

Copyright © 2019 McComb et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Elias McComb, emccomb@som.umaryland.edu.

Received 20 November 2018

Accepted 7 January 2019

Published 31 January 2019

genes (9). A single clustered regularly interspaced short palindromic repeat (CRISPR) locus was detected using CASC v2.5 (10).

Data availability. The whole-genome sequence of *L. crispatus* CO3MRS11 has been deposited in GenBank under the accession no. [CP033426](https://ncbi.nlm.nih.gov/nuccore/CP033426). Raw sequencing reads of the genome of *L. crispatus* CO3MRS11 are available in the SRA under the accession no. [PRJNA499123](https://www.ncbi.nlm.nih.gov/sra/PRJNA499123). Marker gene phylogeny and gene annotation were deposited in Kbase (<https://narrative.kbase.us/narrative/ws.37413.obj.19>) for public assessment.

ACKNOWLEDGMENTS

Research reported in this publication was supported in part by the National Institute of Allergy and Infectious Diseases (NIAID) and the National Institute of Nursing Research (NINR) of the National Institutes of Health under award no. U19AI084044, R01AI116799, and R01NR015495. J. Holm was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award no. F32AI136400.

The institutional review board at the University of Maryland School of Medicine approved the protocol for participant sample collection. The university's guidelines were followed in the conduct of the clinical research. This study was registered at clinicaltrials.gov under the identification no. NCT00576797.

REFERENCES

- Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, Brotman RM, Davis CC, Ault K, Peralta L, Forney LJ. 2011. Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* 108(Suppl 1):4680–4687. <https://doi.org/10.1073/pnas.1002611107>.
- 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. <https://doi.org/10.1128/JB.00399-10>.
- Verstraelen H, Verhelst R, Claeys G, De Backer E, Temmerman M, Vanechoutte M. 2009. Longitudinal analysis of the vaginal microflora in pregnancy suggests that *L. crispatus* promotes the stability of the normal vaginal microflora and that *L. gasseri* and/or *L. iners* are more conducive to the occurrence of abnormal vaginal microflora. *BMC Microbiol* 9:116. <https://doi.org/10.1186/1471-2180-9-116>.
- Borgdorff H, Gautam R, Armstrong SD, Xia D, Ndayisaba GF, van Teijlingen NH, Geijtenbeek TB, Wastling JM, van de Wijkert JH. 2016. Cervico-vaginal microbiome dysbiosis is associated with proteome changes related to alterations of the cervicovaginal mucosal barrier. *Mucosal Immunol* 9:621–633. <https://doi.org/10.1038/mi.2015.86>.
- Johnson B, Selle K, O'Flaherty S, Goh YJ, Klaenhammer T. 2013. Identification of extracellular surface-layer associated proteins in *Lactobacillus acidophilus* NCFM. *Microbiology* 159:2269–2282. <https://doi.org/10.1099/mic.0.070755-0>.
- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
- Price MN, Dehal PS, Arkin AP. 2010. FastTree 2: approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
- Bland C, Ramsey TL, Sabree F, Lowe M, Brown K, Kyrpidis NC, Hugenholtz P. 2007. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics* 8:209. <https://doi.org/10.1186/1471-2105-8-209>.