Complete Genome Sequence of *Lactobacillus crispatus* AB70, Isolated from a Vaginal Swab from a Healthy Pregnant Korean Woman

Dong-Ho Chang,a,b Moon-Soo Rhee,c Sung-Ki Lee,d In-Hwa Chung,a Haeyoung Jeong,f,g Byoung-Chan Kim,a,h

a Metabolic Regulation Research Center, Korea Research Institute of Bioscience and Biotechnology (KRBIBB), Daejeon, Republic of Korea  
b Department of Biological Science and Biotechnology, Hannam University, Daejeon, Republic of Korea  
c Korean Collection for Type Cultures (KCTC), Korea Research Institute of Bioscience and Biotechnology, Jeongeup-si, Jeollabuk-do, Republic of Korea  
d Department of Obstetrics and Gynecology, College of Medicine, Myunggok Medical Research Center, Konyang University, Daejeon, Republic of Korea  
e Sungwun Pharmacopia Co., Ltd., Seongnam-si, Gyeonggi-do, Republic of Korea  
f Infectious Disease Research Center, KRBIBB, Daejeon, Republic of Korea  
g Department of Biosystems and Bioengineering, KRBIB School of Biotechnology, Korea University of Science and Technology (UST), Daejeon, Republic of Korea  
h Department of Bioprocess Engineering, KRBIB School of Biotechnology, UST, Daejeon, Republic of Korea

**ABSTRACT** The vaginal bacterial strain AB70, belonging to the species *Lactobacillus crispatus*, was isolated from a vaginal swab from a healthy pregnant Korean woman. Here, we report the 2.37-Mb complete genome sequence of this strain.

The vaginal microbiota is known to play a significant role in women’s health and pregnancy (1, 2). Several studies have shown that the vaginal microbiome of healthy pregnant women is dominated by *Lactobacillus* spp. and is characterized by less richness and diversity than the microbiome of nonpregnant women (3). A poor dominance of vaginal lactobacilli may cause microbial imbalance in the vagina, often resulting in bacterial vaginosis and high-risk pregnancies (4).

To investigate vaginal microbiota using both culture-dependent and culture-independent techniques, vaginal swabs from healthy pregnant Korean women were collected under anaerobic conditions, as previously described (5). This study received ethical approval from the Konyang University Hospital Institutional Review Board (IRB) (approval number 2014-06-009). Using anaerobic de Man, Rogosa, and Sharpe (MRS) agar plates, several hundred purified vaginal microbes were isolated to purity using single-colony isolation, and most were assigned within *Lactobacillus crispatus* based on 16S rRNA sequences (100% 16S rRNA gene sequence similarities to that of *L. crispatus* strain ATCC 43058T). The 16S rRNA gene was amplified by colony PCR with the universal PCR primers 27F and 1492R (6), and purified PCR products were sequenced by the BIOFACT Co., Ltd. (Daejeon, Republic of Korea).

Antibacterial activities of ethyl acetate extracts were evaluated with disk diffusion assays against *Escherichia coli* KCTC 2441 and *Bacillus cereus* KCTC 3624. One of the isolates, named strain AB70, displayed strong antibacterial activities and was selected for further investigation, including complete genome analysis.

The genomic DNA of strain AB70 was extracted from 1-liter anaerobic MRS liquid cultures, as previously described (5). Genomic library construction and sequencing were carried out at the Chun Lab (Seoul, Republic of Korea) using the PacBio RS II platform (Pacific Biosciences, Menlo Park, CA, USA) with P6-C4 chemistry. We produced 150,292 sequence reads (ca. 1.1 Gbp) from one single-molecule real-time sequencing (SMRT) cell, and these were filtered with default parameters (read quality, ≥0.75; read length, ≥50 bp) with SMRT Analysis version 2.3.0. The resulting 111,972 filtered reads (960.4 Mb total; average length, 8,576 bp) were assembled into two contigs with Canu.
(7) version 1.1 (parameters: genomeSize = 2.4 m, pacbio-raw) and circularized with Circlator (8) version 1.5.1. Residual errors were corrected by running two consecutive rounds of the RS_Resequencing.1 protocol in SMRT Analysis. Genome annotation was carried out with the NCBI Prokaryotic Genome Annotation Pipeline (9).

*Lactobacillus crispatus* AB70 has a chromosome (2,351,263 bp, 37.3% G+C content) that encodes 2,333 protein-coding sequences, 71 tRNAs, and 5 tRNA gene clusters, while the circular plasmid pLcAB70 (16,662 bp, 35.8% G+C content) encodes only 19 protein-coding sequences. The average nucleotide identity between AB70 and the type strain ATCC 33820 calculated with the OrthoANIu algorithm (10) was 97.5%. Genome coverages, however, were only 50.7% (AB70) and 60.0% (ATCC 33820), in accordance with the considerable genetic diversity between strains (11). The genome most similar to AB70 among all 57 publicly available *L. crispatus* genomes in the RefSeq database, identified with dRep (12) version 2.0.5, was that of strain UMB0803 (GenBank number GCF_002861765.1, 99.7% average nucleotide identity [ANI]), one of the nine *L. crispatus* strains isolated from female urine collected with a transurethral catheter (BioProject number PRJNA316969, 200 strains).

The *L. crispatus* AB70 genome contains no virulence-related or resistance-related genes, as assessed with IslandViewer 4 (13), which identified 266,727-bp putative genomic islands (11.3% of the chromosome), distributed into 17 regions. We also identified two bacteriocin biosynthetic gene clusters with antiSMASH version 4.1.0 (14), one of which was similar to a gassericin T (15) biosynthetic gene cluster.

The genomic information from strain AB70 will improve our understanding of the genetic diversity of the species *L. crispatus*, a major member of the vaginal microbiome of healthy pregnant women.

**Data availability.** The complete genome sequences of *L. crispatus* AB70 have been deposited in DDBJ/ENA/GenBank under the accession numbers CP026503 (chromosome) and CP026504 (plasmid). The versions described in this paper are versions CP026503.1 and CP026504.1. Raw sequencing reads are available in NCBI BioProject under the accession number PRJNA431864.

**ACKNOWLEDGMENTS**

This work was supported by the Korea Health Technology R&D Project (HI14C0368) and the KRIBB Research Initiative Program.

**REFERENCES**

1. Kroon SJ, Ravel J, Huston WM. 2018. Cervicovaginal microbiota, women’s health, and reproductive outcomes. Fertil Steril 110:327–336. https://doi.org/10.1016/j.fertstert.2018.06.036.

2. Ollberding NJ, Völgyi E, Macaluso M, Kumar R, Morrow C, Ty lavsky FA, Plyathlakje CJ. 2016. Urinary microbiota associated with preterm birth: results from the Conditions Affecting Neurocognitive Development and Learning in Early Childhood (CANDLE) study. PLoS One 11:e0162302. https://doi.org/10.1371/journal.pone.0162302.

3. Freitas AC, Chaban B, Bocking A, Rocco M, Yang S, Hill JE, Money DM, VOGUE Research Group. 2017. The vaginal microbiome of pregnant women: the circular genome from strain AB70. Sci Rep 7:18818. https://doi.org/10.1038/s41598-017-07790-9.

4. DiGiulio DB, Callahan BJ, McMurdie PJ, Costello EK, Lyell DJ, Robaczewska A, Sun CL, Goltsman DSA, Wong RJ, Shaw G, Stevenson DK, Holmes SP, Relman DA. 2015. Temporal and spatial variation of the human microbiota during pregnancy. Proc Natl Acad Sci U S A 112:11060–11065. https://doi.org/10.1073/pnas.1502875112.

5. Chang D-H, Rhee M-S, Kim B-C. 2016. *Dermabacter vaginatis* sp. nov., isolated from human vaginal fluid. Int J Syst Evol Microbiol 66:1881–1886. https://doi.org/10.1099/ijsem.0.000960.

6. Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR. 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proc Natl Acad Sci U S A 82:6955–6959. https://doi.org/10.1073/pnas.82.20.6955.

7. Koren S, W alenz 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.

8. Hunt M, De Silva N, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. Genome Biol 16:294. https://doi.org/10.1186/s13059-015-0849-0.

9. Tatusova T, DiCuccio M, Badr etdin A, Chetverin M, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.

10. Yoon S-H, Ha S-M, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie van Leeuwenhoek 110:1281–1286. https://doi.org/10.1007/s10482-017-0844-4.

11. Abdelmaksoud AA, Koparde VN, Sheth NU, Serrano MG, Glasscock AL, Fettweis J M, Strauss JF III, Buck GA, Jefferson KK. 2016. Comparison of *Lactobacillus crispatus* isolates from *Lactobacillus*-dominated vaginal microbiomes with isolates from microbiomes containing bacterial vaginosis-associated bacteria. Microbiology 162:466–475. https://doi.org/10.1099/mic.0.002338.

12. Olm MR, Brown CT, Brooks B, Banfield JF. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J 11:2864–2868. https://doi.org/10.1038/ismej.2017.126.

13. Bertelli C, Laird MR, Williams KP, Simon Fraser University Research Comput ing Group, Lau BY, Hoad G, Winsor GL, Brinkman FSL. 2017. IslandViewer 4:
expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res 45:W30–W35. https://doi.org/10.1093/nar/gkx343.

14. Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, Suarez Duran HG, de los Santos ELC, Kim HU, Nave M, Dickschat JS, Mitchell DA, Shelest E, Breitling R, Takano E, Lee SY, Weber T, Medema MH. 2017. antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. Nucleic Acids Res 45:W36–W41. https://doi.org/10.1093/nar/gkx319.

15. Kawai Y, Takahashi R, Arakawa K, Saito T. 2006. Structural and functional characteristics of gassericin A and T produced by Lactobacillus gasseri LA39 and LA158. Jpn J Lactic Acids Bacteria 17:32–39. (In Japanese). https://doi.org/10.4109/jslab.17.32.