



Genome Sequence of *Lactobacillus futsaii* Y97, a Potential Probiotic Strain Isolated from Futsai of Taiwan

Xiaoyu Xu,^a Fang Yao,^a Qiufeng Gou,^a Fuying Dai,^a Qu Pan^a

^aDepartment of Pathogenic Biology, Chengdu Medical College, Sichuan, China

ABSTRACT Here, we report the complete genome sequence of *Lactobacillus futsaii* Y97, a potential probiotic strain isolated from futsai of Taiwan. The genome consists of one chromosome of 2.56 Mb and three plasmids. The genome contains 2,622 genes, which make up 87.06% of the genome.

Lactobacillus futsaii is a Gram-positive rod-shaped nonmotile bacterium belonging to the lactic acid bacteria (1). Traditional fermented foods are rich sources of microorganisms that show probiotic and, hence, health-promoting characteristics (2, 3). *L. futsaii* shows the characteristics of acid resistance, bile salt tolerance, and intestinal colonization. Furthermore, it could be a new strain of probiotic starter for producing gamma-aminobutyric acid (GABA) (4).

L. futsaii Y97 was isolated from futsai, a traditionally home-brewed mustard product produced by Hakka tribes in Taiwan. The genome of *L. futsaii* Y97 was sequenced to identify its specific genetic components and explore its biological characteristics. *L. futsaii* Y97 was grown at 37°C for 48 h under static anaerobic conditions in MRS medium, which was usually used to isolate lactic acid bacteria. The culture was sent to the Beijing Genomics Institute (BGI; Shenzhen, China), and the genomic DNA was extracted using a MiniBEST bacterial genomic DNA extraction kit (TaKaRa, Dalian, China). The 2100 bioanalyzer (Agilent Technologies, Palo Alto, CA) was used to detect the quality of the genomic DNA of *L. futsaii* Y97, and then the genomic DNA was broken to pieces with a desired size by a Covaris S/E210 or g-TUBE instrument. A 270-bp insert library with a read length of 2×150 bp was constructed using a PCR protocol, after the blunt phosphorylated adapters were ligated to the ends of the DNA fragments, and sequenced using the Illumina HiSeq 4000 platform. A 10-kb template library was constructed using a SMRTbell protocol, by which both ends of the DNA fragment were ligated to the connector of the hairpin structure. The 10-kb template library was sequenced using the PacBio Sequel platform. Four single-molecule real-time (SMRT) cell zero-mode waveguide arrays of sequencing were used to generate the subread set. After filtering subreads (length, <1 kb) and adapter sequences, SMRT sequencing generated 4,878,353,482 bp of clean data, with a total of 494,569 subreads and a mean read length of 9,863 bp. The subreads were used for *de novo* assembly with Hierarchical Genome Assembly Process (HGAP) version 3 in SMRT analysis version 2.3.0 software (<https://www.pacb.com/documentation/smrt-analysis-software-installation-v2-3-0/>) (5), which yielded four contigs with an N_{50} value of 11,995 bp. Correction of the PacBio assembly was performed by soapSNP and soapIndel software with default parameters using 851 Mb of clean data from the Illumina HiSeq 4000 sequencing, which were obtained by removal of low-quality reads and adapters and duplication contamination from 1,246 Mb of raw reads (6). The single-base quality of the genome reached 0.9999 after being polished with Quiver. The genome coverage values were $321\times$ with the Illumina HiSeq 4000 platform and $1,841\times$ with the PacBio Sequel platform. Gene prediction was performed with Glimmer version 3.02 (<http://ccb.jhu.edu/software/glimmer/index.shtml>) (7). The functional annotation was ac-

Citation Xu X, Yao F, Gou Q, Dai F, Pan Q. 2019. Genome sequence of *Lactobacillus futsaii* Y97, a potential probiotic strain isolated from futsai of Taiwan. Microbiol Resour Announc 8:e00747-19. <https://doi.org/10.1128/MRA.00747-19>.

Editor Catherine Putonti, Loyola University Chicago

Copyright © 2019 Xu 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 Qu Pan, ppqlove@126.com.

Received 24 June 2019

Accepted 5 September 2019

Published 26 September 2019

complicated by BLAST with the nonredundant (NR) database, Swiss-Prot (<https://www.uniprot.org/uniprot/?query=reviewed:no%20taxonomy:1423818>), Trembl (<https://www.uniprot.org/uniprot/?query=reviewed:no%20taxonomy:1423818>), Antibiotic Resistance Genes Database (ARDB) (<https://card.mcmaster.ca/>), Pathogen Host Interactions (PHI) (<http://www.phi-base.org/>), COG (<https://www.ncbi.nlm.nih.gov/COG/>), Carbohydrate-Active enZymes (CAZy) database (<http://www.cazy.org/>), KEGG database (<http://www.genome.jp/kegg/>), and Gene Ontology (GO) database (<http://geneontology.org/>) (8–14).

The complete genome of *L. futsaii* Y97 consists of a circular chromosome of 2,558,218 bp and three circular plasmids (37,880 bp, 27,087 bp, and 25,584 bp), with G+C contents of 35.68%, 39.44%, 38.93%, and 34.80%, respectively. The genome contains 2,622 genes; the total length of the genes is 2,306,070 bp, which makes up 87.06% of the genome. The number of tandem repeat sequences is 71; the total length of the tandem repeat sequences is 4,575 bp, which makes up 0.1727% of the genome. Also, 43 minisatellite DNAs, 3 microsatellite DNAs, 56 tRNAs, 12 rRNAs, and 7 small RNAs (sRNAs) were predicted. The genome sequence information presented here will help further specific studies of this strain and to exploit its probiotic potential.

Data availability. The genome sequence was deposited in GenBank (BioProject accession number [PRJNA545382](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA545382)) under accession numbers [CP040736](https://www.ncbi.nlm.nih.gov/nuclseq/CP040736), [CP040737](https://www.ncbi.nlm.nih.gov/nuclseq/CP040737), [CP040738](https://www.ncbi.nlm.nih.gov/nuclseq/CP040738), and [CP040739](https://www.ncbi.nlm.nih.gov/nuclseq/CP040739) and SRA accession numbers [SRR9164841](https://www.ncbi.nlm.nih.gov/sra/SRR9164841) and [SRR9157803](https://www.ncbi.nlm.nih.gov/sra/SRR9157803). The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

This research was funded by the National Natural Science Foundation of China (grant number 31170007) and the Health and Family Planning Commission of Sichuan China (grant number 17PJ494).

REFERENCES

- Chao S-H, Kudo Y, Tsai Y-C, Watanabe K. 2012. *Lactobacillus futsaii* sp. nov. isolated from fu-tsai and suan-tsai, traditional Taiwanese fermented mustard products. *Int J Syst Evol Microbiol* 62:489–494. <https://doi.org/10.1099/ijs.0.030619-0>.
- Wang C-Y, Lin P-R, Ng C-C, Shyu Y-T. 2010. Probiotic properties of *Lactobacillus* strains isolated from the feces of breast-fed infants and Taiwanese pickled cabbage. *Anaerobe* 16:578–585. <https://doi.org/10.1016/j.anaerobe.2010.10.003>.
- Oh YJ, Jung DS. 2015. Evaluation of probiotic properties of *Lactobacillus* and *Pediococcus* strains isolated from Omegisool, a traditionally fermented millet alcoholic beverage in Korea. *LWT-Food Sci Technol* 63: 437–444. <https://doi.org/10.1016/j.lwt.2015.03.005>.
- Sanchart C, Rattanaporn O, Haltrich D, Phukpattaranont P, Maneerat S. 2016. Technological and safety properties of newly isolated GABA-producing *Lactobacillus futsaii* strains. *J Appl Microbiol* 121:734. <https://doi.org/10.1111/jam.13168>.
- Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
- Li S, Li R, Li H, Lu JR, Li YR, Bolund L, Schierup MH, Wang J. 2013. SOAPindel: efficient identification of indels from short paired reads. *Genome Res* 23:195–200. <https://doi.org/10.1101/gr.132480.111>.
- Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23: 673–679. <https://doi.org/10.1093/bioinformatics/btm009>.
- UniProt Consortium. 2015. UniProt: a hub for protein information. *Nucleic Acids Res* 43:D204–D212. <https://doi.org/10.1093/nar/gku989>.
- Liu B, Pop M. 2009. ARDB—Antibiotic Resistance Genes Database. *Nucleic Acids Res* 37:D443–D447. <https://doi.org/10.1093/nar/gkn656>.
- Torto-Alalibo T, Collmer CW, Gwinn-Giglio M. 2009. The Plant-Associated Microbe Gene Ontology (PAMGO) Consortium: community development of new Gene Ontology terms describing biological processes involved in microbe-host interactions. *BMC Microbiol* 9:S1. <https://doi.org/10.1186/1471-2180-9-S1-S1>.
- Galperin MY, Makarova KS, Wolf YI, Koonin EV. 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res* 43:D261–D269. <https://doi.org/10.1093/nar/gku1223>.
- Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. 2013. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels* 6:41. <https://doi.org/10.1186/1754-6834-6-41>.
- Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 44:D457–D462. <https://doi.org/10.1093/nar/gkv1070>.
- Ashburner M, Ball CA, Blake JA, Botstein D, Cherry JM. 2000. Gene Ontology: tool for the unification of biology. *Gene* 25:25–29. <https://doi.org/10.1038/75556>.