Draft Genome Sequence of *Lactococcus lactis* subsp. *lactis* W8, a Potential Nisin-Producing Starter Culture for Indian Traditional Fermented Milk (Dahi)

Suranjita Mitra,[a] Bidhan Chandra Mukhopadhyay,[a] Tawsif Ahmed Kazi,[a] Rajarshi Bhattacharya,[a] Sukhendu Mandal,[b] Swadesh Ranjan Biswas[a]

[a]Department of Botany, Visva-Bharati, Santiniketan, West Bengal, India
[b]Department of Microbiology, University of Calcutta, Kolkata, India

ABSTRACT Dahi is a traditional Indian fermented milk consumed regularly as part of the diet because of its palatability and health benefits. Here, we report the draft genome sequence of a unique strain of *Lactococcus lactis* subsp. *lactis*, W8, a lactic acid bacterium that produces nisin while fermenting milk to dahi.

*L*actococcus lactis* is a food-grade Gram-positive lactic acid bacterium widely used in the dairy industry for the production of cheese. Certain *L. lactis* strains are reported to produce nisin, an antibacterial peptide commonly used as a natural food preservative (1). We reported a nisin-producing bacterium from naturally fermented milk (2), which was identified as *L. lactis* subsp. *lactis* W8 by biochemical and 16S rRNA gene sequence analysis (99% identity with *L. lactis* subsp. *lactis*). The uniqueness of this strain lies in its ability to produce nisin while fermenting milk to dahi (pH 4.2) within 6 h (3). Dahi displays antibacterial activity against food spoilage and pathogenic bacteria (3). The strain W8 grows and produces nisin even in 20-times-diluted skim milk (4). Thus, the strain W8 has improved characteristics for the economic production of nisin and fermented milk of safe quality (3). It survives at pH 3.0 and grows in the presence of bile (5). To explore the technological traits relevant to milk fermentation and the probiotic potential, the *L. lactis* W8 genome was sequenced.

Lactococcus lactis W8 was grown overnight at 37°C in tryptone, glucose, and yeast extract (TGE) broth, all at 1% (pH 6.5) (6), and genomic DNA was prepared using Qiagen Genomic-tip 100/G. The paired-end libraries were prepared using an Illumina TruSeq Nano DNA library prep kit and sequenced on a NextSeq 500 instrument with 2 × 150-bp chemistry at Eurofins Genomics India Pvt. Ltd. The data generated were processed to obtain high-quality reads using Trimmomatic v0.35 (7), quality checked using FastQC v0.11.7 (8), and *de novo* assembled using Velvet v1.2.10 (9) with default parameters. The contigs were scaffolded using SSPACE-standard v3.0 (10) with default parameters, generating 25 scaffolds (2,439,471 bp; GC content, 34.9%; N_{50}, 147,833 bp) from 4,116,876 paired-end reads. Gene prediction and functional annotation were done using Blast2GO 5 (11) and the Rapid Annotations using Subsystems Technology (RAST) server (12). A total of 2,444 coding sequences with 58 RNA genes were annotated in the genome of W8. A cluster of 11 genes responsible for nisin production was annotated using BAGEL3 (13). The annotation predicted a number of milk utilization genes encoding *lacG*, *lacZ*, lactose-specific transporters, regulators, several proteinases, including Clp, peptidases, and oligopeptide and amino acid transporters. Unlike the reported dairy *L. lactis* strain (14), the biosynthetic potential of W8 for histidine, leucine, isoleucine, and valine was predicted using BlastKOALA (15), which might facilitate rapid growth in milk deficient in these free amino acids.

The potential of W8 to survive and persist in the gastrointestinal (GI) tract is

predicted by the presence of two cholylglycine hydrolase genes, bile and acid stress resistance genes, and two adhering genes encoding fibronectin/fibrinogen-binding proteins. Further insight into the genome will create a new possibility for the production of novel probiotic dairy products with a built-in preservative.

Data availability. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number QYRO00000000. The version described in this paper is the first version, QYRO01000000. Raw sequencing data have been deposited in the NCBI Sequence Read Archive (16) database under the accession number SRR7868294.

ACKNOWLEDGMENTS

REFERENCES

1. Shin JM, Gwak JW, Kamarajan P, Fenno JC, Rickard AH, Kapila YL. 2016. Biomedical applications of nisin. J Appl Microbiol 120:1449–1465. https://doi.org/10.1111/jam.13033.

2. Mitra S, Chakrabartty PK, Biswas SR. 2005. Production and characterization of nisin-like peptide produced by a strain of *Lactococcus lactis* isolated from fermented milk. Curr Microbiol 51:183–187. https://doi.org/10.1007/s00284-005-4545-2.

3. Mitra S, Chakrabartty PK, Biswas SR. 2010. Potential production and preservation of dahi by *Lactococcus lactis* W8, a nisin producing strain. LWT Food Sci 43:337–342. https://doi.org/10.1016/j.lwt.2009.08.013.

4. Mitra S, Mukhopadhyay BC, Biswas SR. 2009. A foodgrade preparation of nisin in diluted milk by *Lactococcus lactis* W8 and its direct application in milk preservation. J Food Prot 72:2615–2617. https://doi.org/10.4315/0362-028X-72.12.2615.

5. Mitra S, Mukhopadhyay BC, Chakrabartty PK, Biswas SR. 2013. Effect of bile on nisin-mediated antibacterial activity and the expression of nisin genes of *Lactococcus lactis* W8. Curr Microbiol 67:668–673. https://doi.org/10.1007/s00284-013-0417-3.

6. Biswas SR, Ray P, Johnson MC, Ray B. 1991. Influence of growth conditions on the production of bacteriocin, pediocin ACH by *Pediococcus acidilactici* H. Appl Environ Microbiol 57:1265–1267.

7. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

8. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

9. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821–829. https://doi.org/10.1101/gr.074492.107.

10. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27:578–579. https://doi.org/10.1093/bioinformatics/btq683.

11. Conesa A, Götz S, García-Gómez JM, Terol J, Talon M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676. https://doi.org/10.1093/bioinformatics/bti610.

12. 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. https://doi.org/10.1186/1471-2164-9-75.

13. van Heel AJ, de Jong A, Montalbán-López M, Kok J, Kuipers OP. 2013. BAGEL3: automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. Nucleic Acids Res 41:W448–W453. https://doi.org/10.1093/nar/gkt391.

14. Chambellon E, Yvon M. 2003. CodY-regulated aminotransferases AraT and BcaT play a major role in the growth of *Lactococcus lactis* in milk by regulating the intracellular pool of amino acids. Appl Environ Microbiol 69:3061–3068. https://doi.org/10.1128/AEM.69.6.3061-3068.2003.

15. Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol 428:726–731. https://doi.org/10.1016/j.jmb.2015.11.006.

16. Leinonen R, Sugawara H, Shumway M. 2011. The Sequence Read Archive. Nucleic Acids Res 39:D19–D21. https://doi.org/10.1093/nar/gkq1019.