AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequence of a *Lactobacillus fermentum* Strain Isolated from Domestic Sewage in Kerala, India

Pradeesh Babu,[a] Archana Palillam Veedu,[a] Vidhya Prakash,[a] Megha Prasad,[a] Amrita Salim,[a] Ajith Madhavan,[a] Bipin G. Nair,[a] Sanjay Pal[a]

[a]School of Biotechnology, Amrita Vishwa Vidyapeetham, Kollam, Kerala, India

Pradeesh Babu and Archana Palillam Veedu contributed equally to this work. Author order was determined by the order of seniority.

ABSTRACT We report the draft genome sequence of a putative probiotic strain, *Lactobacillus fermentum* ASBT-2, isolated from domestic sewage in Kerala, India. The strain showed probiotic properties (tolerance to low pH and bile salts, binding to host matrix) and reduced the coliform count by 90% in a biofilter used to treat wastewater.

We have developed a microbiome engineering tool to treat wastewater and food with potential probiotic strains and bacteriophages isolated from domestic sewage. *Lactobacillus fermentum* is recognized as a potential probiotic strain with antimicrobial, antioxidative, and cholesterol reduction properties (1–5). The organism was isolated from domestic sewage in Kerala, India, cultured in selective medium, De Man, Rogosa and Sharpe agar (MRS) (6), and confirmed with 16S rRNA gene ribotyping (7).

Genomic DNA was extracted using the phenol-chloroform method (8). The paired-end sequencing library was prepared using the TruSeq Nano DNA library prep kit with an average library size of 478 bp. The Illumina HiSeq platform was used for sequencing (9, 10) the paired-end library, with a read length of 2 × 150 bp. Both quantity and quality checks of the amplified library were performed in a Bioanalyzer 2100 instrument (Agilent Technologies) using a high-sensitivity DNA chip per the manufacturer's instructions. High-quality (5.63 Gb) data, obtained after filtering the reads through Trimmomatic (v0.30) with a quality value (QV) of >20, were used for assembly. All the software settings used were under the default parameters unless otherwise mentioned. *De novo* assembly of paired-end reads was performed using Velvet v1.2.10. (11), and assembly was optimized with a kmer value of 121. The gaps of the assembled scaffold were filled using Gapcloser v1.12 (12). The total number of reads was 37,902,034, and the details of the assembled genome are listed in Table 1. tRNAscan-SE v1.3.1 was used for identification of probable tRNA genes (13). RNAmmer v1.2 was used for rRNA gene identification (14), which yielded a total of five 5S rRNAs and one 16S rRNA.

The 64 scaffolds obtained from *de novo* assembly were subjected to gene prediction using Prodigal v2.6.3 (15), which resulted in the identification of 2,019 coding sequences. The predicted proteins of genes were subjected to a similarity search against NCBI's nonredundant (nr) database using the BLASTP algorithm. Out of 2,019 predicted proteins, 1,989 got a hit in the NCBI database; the remaining 30 were novel proteins. Simultaneously, all the 2,019 proteins were searched for similarity against the UniProt, COG, and Pfam databases using BLASTP with an E value threshold of $1e^{-5}$.

Citation Babu P, Veedu AP, Prakash V, Prasad M, Salim A, Madhavan A, Nair BG, Pal S. 2020. Draft genome sequence of a *Lactobacillus fermentum* strain isolated from domestic sewage in Kerala, India. Microbiol Resour Announc 9:e00713-20. https://doi.org/10.1128/MRA.00713-20.

Editor Julie C. Dunning Hotopp, University of Maryland School of Medicine

Address correspondence to Sanjay Pal, spal05@gmail.com.

Received 18 June 2020
Accepted 29 June 2020
Published 16 July 2020

TABLE 1 Summary of the sequencing details of ASBT-2

Characteristic	Value
Total no. of contigs	74
Total no. of scaffolds	64
Total genome size, including gaps (Ns) (bp)	2,028,463
Total genome size, without gaps (Ns) (bp)	2,027,369
Contig N_{50} (bp)	70,502
Scaffold N_{50} (bp)	70,502
Avg scaffold length (bp)	31,695
Maximum scaffold length (bp)	146,327
GC content (%)	51.89
No. of tRNAs decoding standard 20 amino acids	54
No. of 5S rRNAs	5
No. of 16S rRNAs	1
Total no. of protein-coding genes	2,019
Total gene length (bp)	1,748,088
Maximum gene length (bp)	4,473
Avg gene size (bp)	865

Data availability. This whole-genome shotgun project has been deposited at GenBank under BioProject number PRJNA639667, SRA accession number SRR12020697, and BioSample accession number SAMN15244744.

ACKNOWLEDGMENTS

This work was supported by the Reinvent the Toilet Challenge (RTTC) award (2014) funded by the Bill & Melinda Gates Foundation-BIRAC (Government of India) (grant numbers BIRAC/GCI/0067/02/13-RTTC and OPP1107707). We thank Amrita Vishwa Vidyapeetham, University Grant Commission (UGC), for providing funding to roll number 302092 (Pradeesh Babu) through beneficiary code BININ00345622U and the Council of Scientific and Industrial Research (CSIR) for funding provided to Amrita Salim (roll number 318345).

We declare no conflict of interest.

REFERENCES

1. Fuochi V, Volti G, Furneri P. 2017. Probiotic properties of Lactobacillus fermentum strains isolated from human oral samples and description of their antibacterial activity. Curr Pharm Biotechnol 18:139–149. https://doi.org/10.2174/1389201017666161229153530.
2. Mikelsaar M, Sepp E, Štšepetova J, Hütt P, Zilmer K, Kullisaar T, Zilmer M. 2015. Regulation of plasma lipid profile by lactobacillus fermentum (probiotic strain ME-3. DSM14241) in a randomised controlled trial of clinically healthy adults. BMC Nutr 1:27. https://doi.org/10.1186/s40795-015-0020-z.
3. Mikelsaar M, Zilmer M. 2009. Lactobacillus fermentum ME-3: an antimicrobial and antioxidative probiotic. Microb Ecol Health Dis 21:1–27. https://doi.org/10.1080/08910600902815561.
4. Kang M-S, Lim H-S, Oh J-S, Lim Y-J, Wuertz-Kozak K, Harro JM, Shirtliff ME, Achermann Y. 2017. Antimicrobial activity of Lactobacillus salivarius and Lactobacillus fermentum against Staphylococcus aureus. Pathog Dis 75. https://doi.org/10.1093/femspd/ftx009.
5. Varma P, Nisha N, Dinesh KR, Kumar AV, Biswas R. 2011. Anti-infective properties of Lactobacillus fermentum against Staphylococcus aureus and Pseudomonas aeruginosa. J Mol Microbiol Biotechnol 20:137–143. https://doi.org/10.1159/000328512.
6. Reuter G. 1985. Elective and selective media for lactic acid bacteria. Int J Food Microbiol 2:55–68. https://doi.org/10.1016/0168-1605(85)90057-1.
7. Marín M, Garcia-Lechuz JM, Alonso P, Villanueva M, Alcalá L, Gimeno M, Cercenado E, Sánchez-Somolinos M, Radice C, Bouza E. 2012. Role of universal 16S rRNA gene PCR and sequencing in diagnosis of prosthetic joint infection. J Clin Microbiol 50:583–589. https://doi.org/10.1128/JCM.00170-11.
8. Salim A, Babu P, Mohan K, Moorthy M, Raj D, Kallampillil Thirumeni S, Suresh S, Madhavan A, Nair BG, Chattopadhyay S, Pal S. 2019. Draft genome sequence of an Escherichia coli sequence type 155 strain isolated from sewage in Kerala, India. Microbiol Resour Announc 8:e01707-18. https://doi.org/10.1128/MRA.01707-18.
9. Minoche AE, Dohm JC, Himmelbauer H. 2011. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. Genome Biol 12:R112. https://doi.org/10.1186/gb-2011-12-11-r112.
10. Liu L, Hu N, Wang B, Chen M, Wang J, Tian Z, He Y, Lin D. 2011. A brief utilization report on the Illumina HiSeq 2000 sequencer. Mycology 2:169–191.
11. 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.
12. Xu G-C, Xu T-J, Zhu R, Zhang Y, Li S-Q, Wang H-W, Li J-T. 2019. LR_Gapcloser: a tiling path-based gap closer that uses long reads to complete genome assembly. GigaScience 8:giy157. https://doi.org/10.1093/gigascience/giy157.
13. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964. https://doi.org/10.1093/nar/25.5.955.
14. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160.
15. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119.