Acta Scientiarum

http://periodicos.uem.br/ojs/
ISSN on-line: 1807-863X
Doi: 10.4025/actascibiolsci.v45i1.62205

Genome-sequence, annotation and phylogenetic insights of the lactic acid bacterium *Limosilactobacillus fermentum* strain LAB1 obtained from the dairy beverage borhani

Tanim Jabid Hossain[1,2]

[1]Department of Biochemistry and Molecular Biology, University of Chittagong, Chattogram 4331, Bangladesh. [2]Microbial Biochemistry Informatics and Genomics Laboratory, BPM Research Group, Chattogram 4331, Bangladesh. E-mail: tanim.bmb@gmail.com

ABSTRACT. *Limosilactobacillus fermentum* is a promising probiotic with several documented health benefits. LAB1 is an antagonistic *L. fermentum* strain isolated from borhani, a traditional South Asian beverage prepared from dairy and plant ingredients. Here, I present the genome sequence of the *L. fermentum* LAB1 strain, its annotation, and phylogenetic features. The 2.01 Mb genome with a G+C content of 51.9% was assembled into 221 contigs and predicted to have 1,913 protein-coding genes, 98 pseudo genes, 7 rRNAs, 60 tRNAs, and 1 CRISPR array. As much as 91.1% of the coding sequences could be assigned to known functional genes. Determination of average nucleotide identity (ANI) of the genome sequence revealed 99.37% identity to that of the type strain ATCC 14931. Its 16S rRNA gene sequence extracted from the genome sequence showed close phylogenetic association with several *L. fermentum* strains. The genome sequence is expected to provide useful insights with regard to the phenotypic, metabolic and beneficial aspects of this lactic acid bacterium.

Keywords: genome sequence; *Limosilactobacillus fermentum*; lactic acid bacteria; dairy beverage; borhani; phylogenetic association.

Received on January 23, 2022.
Accepted on December 21, 2022.

Introduction

Limosilactobacillus fermentum (formerly known as *Lactobacillus fermentum*) is a Gram-positive rod-shaped lactic acid bacterium (LAB) that has attracted increasing attention in recent years for its promising probiotic potential (Naghmouchi et al., 2020). The bacterium is heterofermentative (Ibrahim, 2016), a producer of important metabolites such as exopolysaccharides (EPS) and biosurfactants, and tolerant to harsh gastrointestinal conditions such as high levels of pH and bile salts (Klayraung, & Okonogi, 2009; Satpute et al., 2016; Naghmouchi et al., 2020). Previous studies reported identification of *L. fermentum* strains from milk, gastrointestinal tract (GIT) and urogenital tract of human, GIT of other animals, plant tissues, and food products (Sepp et al., 2018; Santos et al., 2021). Some strains have been implicated with various health benefits that are strain-specific and include antimicrobial, antioxidative, anti-inflammatory, anti-aging, anti-cancer, immunomodulatory, and cholesterol-lowering effects (Santos et al., 2021). Hence, *L. fermentum* holds high promise for future development as a probiotic.

The bacterium is already being used as starter culture and food preservative due to its ability to improve sensorial properties of foods such as flavor, texture etc. (Yo, Yong, & Oh, 2021). Moreover, a recent study showed that nutraceutical formulations of *L. fermentum* strains combined with phenolic compounds such as quercetin or resveratrol can be an effective technique to obtain beneficial effects from both the microbial and the plant components (Sampaio et al., 2021). Hence, it might be possible to use *L. fermentum* strains for enhancing the bioavailability and bioactivity of (poly)phenols that are otherwise poorly absorbed. In another study, a *L. fermentum* strain isolated from the Chinese fermented food Jiangshui was demonstrated to have a high degradation capacity of uric acid which was further validated in mice model wherein the *L. fermentum* strain could also decrease hyperuricemia (Wu et al., 2021). Another strain also isolated from Jiangshui was able to reduce arsenic accumulation in crayfish (Han et al., 2022). Therefore, *L. fermentum* strains are considered to be strong probiotic candidates for health and food related applications. A further advantage of *L. fermentum* is that this species is already regarded as 'generally recognized as safe' (GRAS) by the US Food and

Drug Administration (FDA) (García et al., 2017). However, since the health promoting effects of lactic acid bacteria are strain-specific wherein a particular strain might only provide one or a few of the beneficial effects, a continuous need exists for the isolation and characterization of new strains with superior natural qualities.

The *L. fermentum* LAB1 strain reported here was isolated from borhani (Hossain, Mozumder, Ali, & Akther, 2022a) which is a heavily spiced dairy beverage prepared from ingredients of both animal and plant origin. Its principal constituents are yoghurt, coriander and mint, and small amounts of pepper, mustard, cumin seeds, green chilies, rock salt, and sugar. Popular among people of all ages, borhani is usually consumed after a heavy meal and believed to be highly beneficial for digestion. Although this traditional South Asian beverage has been locally prepared for many years at home and also served at restaurants, in recent years it has been increasingly produced commercially, and marketed under several renowned or unrecognized brands. The borhani-derived strain of the present study had strong antimicrobial activity against several Gram-positive and Gram-negative pathogenic organisms (Hossain, et al., 2022a), and produced EPS and biosurfactant.

Genome sequencing of probiotic strains can be a valuable genetic resource providing important new insights into the strains' probiotic attributes. Here, I report the genome sequence of the LAB1 strain which is expected to offer useful genetic information into its beneficial properties by the identification and functional characterization of the encoded genes. The taxonomic and phylogenetic association of the strain s also discussed.

Material and methods

Isolation source and maintenance of the LAB1 strain

The LAB1 strain was isolated by the conventional spreading technique (Hossain, Alam, & Sikdar, 2011). Five different samples of borhani purchased from shops in the Chattogram city were mixed, 5× serial dilutions of the mixture were spread on De Man, Rogosa and Sharpe (MRS) agar medium (20 g L^{-1} glucose, 10 g L^{-1} peptone, 10 g L^{-1} beef extract, 5 g L^{-1} yeast extract, 2 g L^{-1} dipotassium hydrogen phosphate, 5 g L^{-1} sodium acetate trihydrate, 2 g L^{-1} triammonium citrate, 0.1 g L^{-1} magnesium sulphate heptahydrate, 0.05 g L^{-1} manganese sulphate tetrahydrate, 1 ml L^{-1} Tween-80, 18 g L^{-1} agar and water) (Microbial Culture Media, 2020) and incubated at 37°C for 48 hours (Hossain, et al., 2022b). Pure cultures were obtained by repeated streaking on MRS medium and preserved in 20% v v^{-1} glycerol at -80°C (Hossain et al., 2020). For subsequent analysis, cells were first activated from the stock (Hossain, Das, Ali, Chowdhury, & Zedny, 2021) in MRS broth which was further transferred and grown in fresh medium.

Extraction of genomic DNA

For whole genome sequencing, activated culture was streaked on MRS agar medium and incubated at 37°C for 48 hours. Subsequently, one of the colonies grown on the agar-plate was used to extract its total genomic DNA using the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA) and quantified using the Quantus fluorometer (Promega) according to the manufacturer's instructions.

Library preparation, sequencing and assembly

Libraries were prepared using the Illumina DNA prep library preparation kit (San Diego, USA) and sequencing was conducted using the Illumina NextSeq 2000 system at Invent Technologies, Bangladesh. For subsequent analysis, default parameters were used for all software unless otherwise specified. The raw reads were quality checked with FastQC (2020) version 0.11.8, filtered for the low-quality reads and reads shorter than 200 bp by Trimmomatic version 0.38 (Bolger, Lohse, & Usadel, 2014), and assembled by SPAdes version 3.12.0 (Bankevich et al., 2012)).

Analysis and annotation of genome sequence

Functional annotations of the genome were carried out by the NCBI prokaryotic genome annotation pipeline (PGAP) (Tatusova et al., 2016) and Rapid Annotation using Subsystem Technology (RAST) (Overbeek et al., 2014). Genome assembly and annotation completeness was assessed by Benchmarking Universal Single-Copy Orthologs (BUSCO) (Simão, Waterhouse, Ioannidis, Kriventseva, & Zdobnov, 2015). Taxonomic annotations were based on average nucleotide identity (ANI) conducted using ANI calculator (http://enve-omics.ce.gatech.edu/ani/index). The taxonomic assignment was further verified by NCBI blastn and Unipept using the sequences (Mesuere et al., 2015).

Genome project

The whole genome project has been deposited in GenBank under the accession number JAJTII000000000.

Taxonomic analysis

Phylotype of the isolate was determined based on sequence similarity of its 16S rRNA gene using NCBI BLAST Suite and RDP classifier as previously described (Hossain et al., 2018).

Phylogenetic analysis

Phylogenetic analysis was performed as described previously (Ali et al., 2021). Briefly, multiple sequence alignment was performed using ClustalW and a tree of the aligned sequences was built by maximum likelihood method using the MEGA software version X (Kumar, Stecher, Li, Knyaz, & Tamura, 2018). Sequences for construction of the phylogenetic tree was retrieved from the GenBank and EZbioCloud databases (Yoon et al., 2017).

Results and discussions

Genome features of *Limosilactobacillus fermentum* LAB1

The draft genome was 2,011,628 nucleotides and consisted of 221 contigs with a G+C content of 51.9% and an N50 value of 34,128 (Table 1). A total of 2,081 genes were predicted by PGAP which included 1,913 protein coding genes, 98 pseudogenes, seven rRNAs, 60 tRNAs, and three ncRNAs. As much as 91.1% of the protein coding sequences was assigned to known functions and the rest was annotated as hypothetical proteins. Analysis with BUSCO revealed a completeness score of 99.2% with a total of 122 complete single-copy genes, one complete and duplicated gene, and one fragmented gene.

Table 1. Genome features of the *Limosilactobacillus fermentum* strain LAB1.

Features	Value
Genome size (bp)	2,011,628
No. of contigs	221
G+C content (%)	51.9
N50 (bp)	34,128
Completeness score (%)	99.2
Genes (total)	2,081
CDSs (total)	2,011
Genes (coding)	1,913
Genes (RNA)	70
rRNAs	2, 2, 3 (5S, 16S, 23S)
tRNAs	60
ncRNAs	3
Pseudo Genes	98
CRISPR array	1

These genome features of LAB1 appear to be similar to those reported for other *L. fermentum* strains. For example, genome of the *L. fermentum* ATCC 23271 strain had 2,193,335 bp with 50.9% GC content and 2,123 protein-coding sequences (Santos et al., 2021). *Limosilactobacillus fermentum* KUB-D18 genome was reported to be approximately 2.02 Mbps having GC content of 51.72% and 2,158 protein-encoding genes (Phujumpa et al., 2022). Strain DM075 had 2,204,022-bp genome with 51.0% GC content and 2,129 coding sequences (Park et al., 2022). According to the NCBI genome database, the 156 *L. fermentum* genome assemblies have a median total length of 2.01183 Mb with a median GC of 51.8% and median protein count of 1,892. For *L. fermentum* reference genome, ASM2281924v1, the genome size, GC content and protein count are 2.02 Mb, 51.8% and 1,888 respectively.

Genetic identification and phylogenetic features

Analysis of both the genomic and 16S rRNA gene sequences supported placement of the LAB1 strain in *Limosilactobacillus* group. The 16S rRNA gene was extracted from its genome sequence predicted by RAST. The sequence identity and sequence-based taxonomy were analyzed by NCBI nucleotide BLAST and RDP classifier. 16S rRNA gene of LAB1 showed a high homology to those in the public databases with 100% identity

to that of *L. fermentum* strain HPLD and more than 99% identity to several other strains of *L. fermentum* with 99 to 100% coverage (Table 1). Phylogenetic tree also showed a close association of the LAB1 strain with *L. fermentum* strains (Figure 1) particularly HPLD and L18, providing further support for its taxonomic affiliation.

Table 2. Molecular identification of the isolate based on homology of its 16S rRNA gene sequence. Top five strains that appeared in the search results are shown.

	NCBI BLAST[1]							RDPI SeqMatch			
SN	L. f.[2] strain	Accession	Score	Query Cover	E value	% ident	Acc. Len	L. f.[2] strain	ID	Score	UCO[3]
1	HPLD	MN336188	2368	1	0	100	1582	UICT/BK1; FJ463754	S001243173	1	0839
2	L18	MK033872	2362	0.99	0	100	1573	AB5-18; FJ966275	S001549769	1	0664
3	K3	EU621850	2361	0.99	0	100	1569	AB21-107; FJ966280	S001549774	1	0673
4	K1	EU621848	2355	0.99	0	99.92	1568	AK2-8; FJ966281	S001549775	1	0581
5	NRIC 0145	AB362626	2355	0.99	0	99.92	1570	AK4-180; FJ966283	S001549777	1	0563

[1]Results for only 16S rRNA gene sequences are included. [2]L. f.: *Limosilactobacillus fermentum*. [3]UCO: Unique common oligomers.

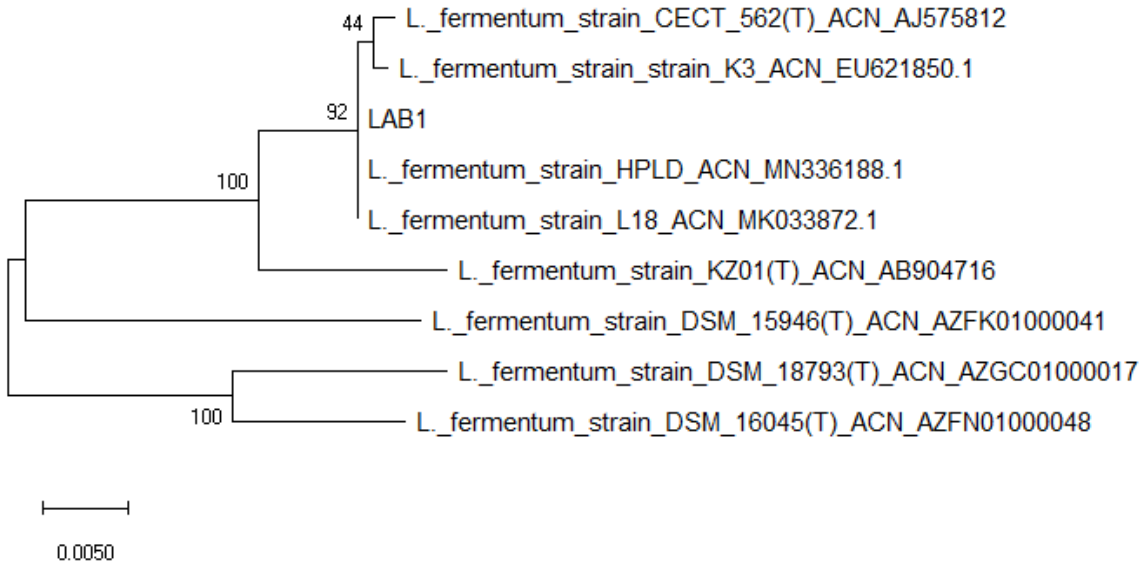

Figure 1. Phylogenetic tree showing evolutionary relationship of the LAB1 strain with other *Limosilactobacillus fermentum* strains including several type strains. The evolutionary history was inferred using Maximum Likelihood method and Tamura-Nei model. Initial tree for the heuristic search was automatically obtained by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site; (T): type strain; ACN: accession number.

ANI values indicated over 99% similarity with the genomes of several *L. fermentum* strains; for example, 99.39% identity with *L. fermentum* ATCC 23271 and 99.37% identity with the type strain *L. fermentum* ATCC 14931, which provides evidence for LAB1 to be classified as a *L. fermentum* strain (Taxonomy ID: 1613).

Conclusion

Genome sequence and annotation reported in this study provides an important genetic resource of the *L. fermentum* strain. The genome information will facilitate genetic characterization of its beneficial and safety properties and their biochemical validation for its development and application as a probiotic. An extensive genome mining is, therefore, required to identify its genes involved in the principal metabolic pathways and biomolecules required for its probiotic functions. Moreover, comparative and functional genome analysis will also provide genetic basis for both the strain-specific and shared beneficial traits present in this species.

Acknowledgements

This work was financially supported by University of Chittagong. The author gratefully acknowledges the support from Nabila Rahman; and also thanks all members of the Biochemistry and Pathogenesis of Microbes–BPM Research Group, Chattogram 4331, Bangladesh.

References

Ali, F., Das, S., Hossain, T. J., Chowdhury, S. I., Zedny, S. A., Das, T. ... Uddin, M. S. (2021). Production optimization, stability and oil emulsifying potential of biosurfactants from selected bacteria isolated from oil-contaminated sites. *Royal Society Open Science, 8*(10). DOI: https://doi.org/10.1098/rsos.211003

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., ... Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology, 19*(5), 455-477. DOI: https://doi.org/10.1089/cmb.2012.0021

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics, 30*(15), 2114-2120. DOI: https://doi.org/10.1093/bioinformatics/btu170

FastQC (2020). *Babraham Bioinformatics*. Retrieved from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

García, A., Navarro, K., Sanhueza, E., Pineda, S., Pastene, E., Quezada, M., ... González, C. (2017). Characterization of *Lactobacillus fermentum* UCO-979C, a probiotic strain with a potent anti-Helicobacter pylori activity. *Electronic Journal of Biotechnology*, *25*, 75-83. DOI: https://doi.org/10.1016/j.ejbt.2016.11.008

Han, R., Khan, A., Linlg, Z., Wu, Y., Feng, P., Zhou, T., ... Liu, P. (2022). Feed-additive *Limosilactobacillus fermentum* GR-3 reduces arsenic accumulation in *Procambarus clarkii*. *Ecotoxicology and Environmental Safety, 231*, 113216. DOI: https://doi.org/10.1016/j.ecoenv.2022.113216

Hossain, T. J., Alam, M. K., & Sikdar, D. (2011). Chemical and microbiological quality assessment of raw and processed liquid market milks of Bangladesh. *Continental Journal of Food Science and Technology, 5*(2), 6-17. DOI: https://doi.org/10.5281/zenodo.5568945

Hossain, T. J., Mozumder, H. A., Ali, F., & Akther, K. (2022a). Inhibition of pathogenic microbes by the lactic acid bacteria *Limosilactobacillus fermentum* Strain LAB-1 and *Levilactobacillus brevis* Strain LAB-5 isolated from the dairy beverage Borhani. *Current Research in Nutrition and Food Science Journal, 10*(3), 928-939. DOI: https://doi.org/10.12944/CRNFSJ.10.3.10

Hossain, T. J., Das, M., Ali, F., Chowdhury, S. I., & Zedny, S. A. (2021). Substrate preferences, phylogenetic and biochemical properties of proteolytic bacteria present in the digestive tract of Nile tilapia (*Oreochromis niloticus*). *AIMS Microbiology, 7*(4), 528-545. DOI: https://doi.org/10.3934/microbiol.2021032

Hossain, T. J., Chowdhury, S. I., Mozumder, H. A., Chowdhury, M. N. A., Ali, F., Rahman, N., & Dey, S. (2020). Hydrolytic exoenzymes produced by bacteria isolated and identified from the gastrointestinal tract of Bombay duck. *Frontiers in Microbiology, 11*. DOI: https://doi.org/10.3389/fmicb.2020.02097

Hossain, T. J., Nafiz, I. H., Ali, F., Mazumder, H. A., Islam, S., Rahman, N., & Khan, S. (2022b). Antipathogenic action and antibiotic sensitivity pattern of the borhani-associated lactic acid bacterium *Weissella confusa* LAB-11 (SSRN Scholarly Paper). Retrieved from https://papers.ssrn.com/abstract=4161742

Hossain, T. J., Manabe, S., Ito, Y., Iida, T., Kosono, S., Ueda, K., ... Suzuki, T. (2018). Enrichment and characterization of a bacterial mixture capable of utilizing C-mannosyl tryptophan as a carbon source. *Glycoconjugate Journal, 35*(2), 165-176. DOI: https://doi.org/10.1007/s10719-017-9807-2

Ibrahim, S. A. (2016). Lactic acid bacteria: *Lactobacillus* spp.: other species. In *Reference Module in Food Science*. Greensboro, NC: North Carolina A & T. State University. DOI: https://doi.org/10.1016/B978-0-08-100596-5.00857-X

Klayraung, S., & Okonogi, S. (2009). Antibacterial and antioxidant activities of acid and bile resistant strains of *Lactobacillus fermentum* isolated from miang. *Brazilian Journal of Microbiology, 40*(40), 757-766. DOI: https://doi.org/10.1590/S1517-83822009000400005

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution, 35*(6), 1547-1549. DOI: https://doi.org/10.1093/molbev/msy096

Mesuere, B., Debyser, G., Aerts, M., Devreese, B., Vandamme, P., & Dawyndt, P. (2015). The Unipept metaproteomics analysis pipeline. *Proteomics, 15*(8), 1437-1442. DOI: https://doi.org/10.1002/pmic.201400361

Microbial Culture Media. (2020). Bacteriological Growth Media Composition, preparation and preservation of nutritional media for culturing bacteria and fungi (for our lab students). *BPM Research Group*. Retrieved from https://sites.google.com/view/bpm-research-group/news-info/culture-media

Naghmouchi, K., Belguesmia, Y., Bendali, F., Spano, G., Seal, B. S., & Drider, D. (2020*). Lactobacillus fermentum*: a bacterial species with potential for food preservation and biomedical applications. *Critical Reviews in Food Science and Nutrition, 60*(20), 3387-3399. DOI: https://doi.org/10.1080/10408398.2019.1688250

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J. Davis, J. J., Disz, T., ... Stevens, R. (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research, 42***(**D206-214). DOI: https://doi.org/10.1093/nar/gkt1226

Park, D.-Y., Hwang, J., Kim, Y., Kim, Y.-Y., Kim, H.-S., & Hwang, I. (2022). Whole-Genome Sequence of *Limosilactobacillus fermentum* Strain DM075, Isolated from the Human Oral Cavity. *Microbiol Resour Announc* e0081922. DOI: https://doi.org/10.1128/mra.00819-22

Phujumpa, P., Muangham, S., Jatuponwiphat, T., Koffas, M., Nakphaichit, M., & Vongsangnak (2022). Comparative genomics-based probiotic relevance of *Limosilactobacillus fermentum* KUB-D18. *Gene,* 840(146747). DOI: https://doi.org/10.1016/j.gene.2022.146747

Sampaio, K. B., Nascimento, Y. M., Tavares, J. F., Cavalcanti, M. T., Alves, J. L. B., Garcia, E. F., & Souza, E. L. (2021). Development and in vitro evaluation of novel nutraceutical formulations composed of *Limosilactobacillus fermentum*, quercetin and/or resveratrol. *Food Chemistry, 342*, 128264. DOI: https://doi.org/10.1016/j.foodchem.2020.128264

Santos, C. I., Campos, C. D. L., Nunes-Neto, W. R., Carmo, M. S., Nogueira, F. A. B., Ferreira, R. M., ... Monteiro-Neto, V. (2021). Genomic analysis of *Limosilactobacillus fermentum* ATCC 23271, a potential probiotic strain with anti-candida activity. *Journal of Fungi, 7*(10), 794. DOI: https://doi.org/10.3390/jof7100794.

Satpute, S. K., Kulkarni, G. R., Banpurkar, A. G., Banat, I. M., Mone, N. S., Patil, R. H., & Cameotra, S. S. (2016). Biosurfactant/s from *Lactobacilli* species: properties, challenges and potential biomedical applications. *Journal of Basic Microbiology, 56*(11), 1140-1158. DOI: https://doi.org/10.1002/jobm.201600143

Sepp, E., Smidt, I., Stsepetova, J., Roop, T., Hutt, P., Ratsep, M., & Mikelsaar, M. (2018). The effect of *Lactobacillus fermentum* ME-3 on the intestinal microbiota and urine polyamines content: a double-blind placebo-controlled pilot trial. *Journal of Functional Foods, 48*, 430-438. DOI: https://doi.org/10.1016/j.jff.2018.04.053

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics, 31*(19), 3210-3212. DOI: https://doi.org/10.1093/bioinformatics/btv351

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., ... Ostell, J. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Research, 44*(14), 6614-6624. DOI: https://doi.org/10.1093/nar/gkw569

Wu, Y., Ye, Z., Feng, P., Li, R., Chen, X., Tian, X., ... Li, X. (2021). *Limosilactobacillus fermentum* JL-3 isolated from "Jiangshui" ameliorates hyperuricemia by degrading uric acid. *Gut Microbes 13*(1), 1-18. DOI: https://doi.org/10.1080/19490976.2021.1897211

Yoo, H., Yong, C. C., & Oh, S. (2021). Genome analysis of *Limosilactobacillus fermentum* JN2019 applied to tumeric fermentation for animal feed. *Journal of Animal Science and Technology, 63*(5), 1204-1206. DOI: https://doi.org/10.5187/jast.2021.e81

Yoon, S.-H., Ha, S.-M., Kwon, S., Lim, J., Kim, Y., Seo, H., & Chun, J. (2017). Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *International Journal of Systematic and Evolutionary Microbiology, 67*(5), 1613-1617. DOI: https://doi.org/10.1099/ijsem.0.001755