Original Article

JHSMR Journal of Health Science and Medical Research

A Comprehensive Approach to Candidate Probiotic Screening through Nanopore Sequencing and Bioinformatic Analysis

Jirasa Boonsan, M.Sc.[1], Kamonnut Singkhamanan, Ph.D.[1], Thunchanok Yaikhan, Ph.D.[1], Ei Phway Thant, M.Sc.[1], Chanida Sakunrang, B.Sc[2], Monwadee Wonglapsuwan, Ph.D.[2], Komwit Surachat, Ph.D.[1,3]

[1]Department of Biomedical Sciences and Biomedical Engineering, Faculty of Medicine, Prince of Songkla University, Songkhla 90110, Thailand.

[2]Division of Biological Science, Faculty of Science, Prince of Songkla University, Songkhla 90110, Thailand.

[3]Translational Medicine Research Center, Faculty of Medicine, Prince of Songkla University, Songkhla 90110, Thailand.

Received 31 August 2024 ● Revised 5 September 2024 ● Accepted 5 September 2024 ● Published online 12 March 2025

Abstract:

Objective: This study aimed to screen and evaluate probiotic properties in bacterial strains isolated from fermented foods and animal feces, using nanopore sequencing and bioinformatic analysis.

Material and Methods: Thirty bacterial strains were isolated from local foods and animal feces within Songkhla, Thailand. After excluding seven pathogenic strains, the remaining 23 strains underwent Deoxyribonucleic Acid (DNA) extraction and whole-genome sequencing using the MinION™ platform from Oxford Nanopore Technologies (ONT). The genomes were then assembled and annotated using bioinformatics analysis to evaluate their probiotic traits and safety profiles.

Results: Genome assembly statistics revealed considerable variability in genome sizes and contig lengths among the strains, with the N50 values ranging from 326,094 bp to 3,226,988 bp. None of the 23 selected strains contained virulence genes associated with significant health risks; confirming their safety. Comprehensive genetic analysis identified key probiotic genes related to acid stress resistance, adhesion, antioxidant activity, bile resistance, and synthesis of beneficial substances; underscoring their potential efficacy in promoting host health.

Conclusion: Whole-genome sequencing, using ONT, combined with bioinformatic analysis, is an effective approach for identifying probiotic strains and predicting their functional properties. This method provides a comprehensive understanding of their potential health benefits and ensures their safety for consumption.

This paper was from the 3rd Annual Health Research International Conference (AHR-iCON, August 29-30, 2024).
Contact: Komwit Surachat, Ph.D.
Department of Biomedical Sciences and Biomedical Engineering, Faculty of Medicine, Prince of Songkla University, Songkhla 90110, Thailand.
E-mail: komwit.s@psu.ac.th

J Health Sci Med ResRes 2025;43(5):e20251175
doi: 10.31584/jhsmr.20251175
www.jhsmr.org

Keywords: bacteria genome, bioinformatic analysis, gene annotation, nanopore sequencing, probiotic

Introduction

Probiotic microorganisms are widely used globally in human and livestock healthcare in the form of food, beverages, dietary supplements, and disease treatments. The main probiotic groups are *Lactobacillus, Bifidobacterium, Pediococcus*, and *Lactococcus.* Previous studies have found that probiotics have properties that prevent and treat several intestinal diseases such as inflammatory bowel disease (IBD), irritable bowel syndrome (IBS), and gastrointestinal infections[1,2]. Probiotics also aid in weight loss and weight management by reducing body fat and improving metabolism[3,4]. Additionally, they help alleviate symptoms of depression and anxiety, and improve mood[5]. Furthermore, probiotics reduce skin inflammation and acne, assist in maintaining healthy skin, and increasing skin hydration[6,7] Finally, probiotics help lower cholesterol and blood pressure levels, which is beneficial for heart health[8].

Commonly, several studies or routine laboratories have used conventional methods to identify probiotic strains, such as using selective media for cultivation, observing the morphology of colony characteristics or under microscopic, performing identification, including biochemical tests, the VITEK 2 system[9], Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometer (MALDI-TOF MS)[10], and Polymerase Chain Reaction (PCR) identification[11]. To screen probiotic properties, traditional tests or phenotypic identification often assess various characteristics, including tolerance to acids and bile salts, antibacterial activity, resistance to phenol, ability to auto aggregate, and safety assessments such as hemolysis activity. Although these methods are helpful in identifying probiotic strains and features, they are time-consuming, laborious and carry a risk of misidentification.

Currently, nanopore sequencing technology, a high-throughput technique from Oxford Nanopore Technologies (ONT), has been introduced as a new method for identifying bacterial species. The MinION™ is one of the next-generation sequencing platforms launched by ONT that can generate long-reads and provide real-time results[12]. This technology is suitable for whole-genome sequencing because it generates long-read sequences, overcoming the limitations of short-read sequencing, which may not be able to assemble a complete genome. This approach leads to convenient characterization of genomic content, including probiotic genes, antimicrobial resistance genes (ARGs), and virulence genes. Various studies have utilized nanopore sequencing for various applications, including 16S rRNA-targeted sequencing (V3-V4) and shotgun metagenomic sequencing, to analyze the composition of probiotics in additional supplements in food[6]. Due to numerous advantages of the technology, such as rapid and real-time sequencing, portability, and the ability to generate long reads, it facilitates a more accurate and comprehensive analysis of microbial communities and microbial insight genomic characteristics.

In this study, we employed nanopore whole-genome sequencing to identify probiotic species and predict their functional properties through the MinION™ platform. This method provided comprehensive information into probiotic genomes, allowing us to understand and predict probiotic properties. Bioinformatics analysis assists in manipulating and interpreting the sequencing data. Long-read sequences were assembled into a draft genome. Afterward, the analysis involved identifying genes associated with significant probiotic properties, such as acids and bile salts tolerance, antimicrobial activities, adherence abilities, and the synthesis of various types of beneficial substances. Safety-related genes were also assessed to ensure the probiotic strains were safe for consumption. This integrated approach allows for a comprehensive understanding of probiotic strains and their potential health benefits. This research provides safer and more effective probiotic screening methods, advancing

our understanding of probiotics and their impact on human health. Additionally, it helps us to select candidate probiotics for specific applications, ensuring targeted and effective use in both clinical and commercial settings.

Material and Methods

Bacterial isolation, strain identification, and probiotic selection

Thirty bacterial strains were isolated from various local sources in Songkhla, Thailand, including fermented sticky rice, fermented fish, pickled vegetables, and animal feces (chicken and goat). Lactobacillus MRS broth was used to culture probiotics in the samples. All strains were enriched in conditions of 37 °C for 18 hours. After enrichment, bacteria were 10-fold diluted and placed on Man-Rogosa-Sharpe (MRS) agar at 37 °C for 24-72 hours under anaerobic conditions to obtain single colonies. The single colonies were collected and cultured to ensure that there was no contamination. All the cultured isolates in MRS broth were mixed with 20% glycerol and stored at -80 °C as stock cultures[13]. The isolated bacteria, maintained at -80 °C, were then cultured on MRS agar to obtain single colonies for species identification using MALDI-TOF MS with a MALDI Bio-Typer (Bruker Daltonics, Karlsruhe, Germany). The list of probiotics and species identification is provided in Table 1. Among the 30 strains, seven were identified as pathogens and were excluded from the study. The remaining strains, presumed to have probiotic properties, were selected for Deoxyribonucleic Acid (DNA) extraction and whole-genome sequencing using Oxford Nanopore Technologies.

Genomic DNA extraction

Genomic DNA was extracted using the ZymoBIOMICS™ DNA Miniprep Kit (Zymo Research, Seattle, United States) following the manufacturer's instructions[14]. The quality and concentration of the gDNA were assessed using agarose gel electrophoresis (1.0% agarose gel in Tris-borate-EDTA buffer, 80V, 45 min) and a Qubit4 fluorometer (Thermo Fisher Scientific, Inc.). According to the nanopore manufacturer's protocol, the concentration of gDNA needed to be at least 200 ng per sample. The bacterial DNA was stored at -20 °C until further use. The genome sizes of the identified bacteria from MALDI-TOF MS were approximately estimated, and these data were used to calculate the amount of data expected from nanopore sequencing (Supplementary Table 1).

Library preparation, sequencing and bioinformatic analysis

Libraries were prepared using the Rapid Barcoding Kit 24 V14 (SQK-RBK114.24, Oxford Nanopore Technologies, Oxford, UK) for sequencing on R.10.4.1 flow cells[15]. The datasets were generated on a MinION Mk1C (Oxford Nanopore Technologies) device with the super-accuracy base-calling mode selected. All other parameters were set to the default settings.

Genome assembly and annotation

The Bactopia pipeline[16], a comprehensive bioinformatics tool for analyzing bacterial genomes, was used to process all probiotic genomes. All analyses and programs in this study were performed using default parameters. Initially, FastQC was employed for quality control of the sequences[17]. Subsequently, genomes that passed the quality control were assembled using Flye, a tool optimized for long-read data from nanopore sequencing. The assembly quality was then evaluated using QUAST[18], which provided metrics such as N50, total length, and number of contigs. The qualified assembled genomes from this step were used for further downstream analysis.

Safety assessment

For the safety assessment, bioinformatic analysis was applied to evaluate safety profiles. The strains that contain antibiotic resistance and virulence factors were considered pathogenic and excluded from this study.

The ABRicate pipeline (https://github.com/tseemann/abricate) was used to search against Comprehensive Antibiotic Resistance Database (CARD) and the Virulence Factor Database (VFDB) to determine the potential ARGs and virulence contained in all probiotic strains[19,20].

Identification of Probiotic Properties Genes

To identify genes associated with probiotic properties, gene profiles were gathered from the literature to construct a database for identifying probiotic traits. Subsequently, ABRicate was used against the custom database for screening probiotic property genes. The output files from ABRicate were then extracted to identify the probiotic genes.

Results

Summary of sequencing data

The genome assembly statistics and identification results for 23 isolated probiotic strains, including *Lactiplantibacillus, Pediococcus*, *Limosilactobacillus*, *Weissella*, and *Enterococcus*, were provided in Supplementary Table 2. The genome sizes of the strains ranged from 1,845,929 bp (*Pediococcus pentosaceus* P1) to 7,139,934 bp (*Lactiplantibacillus plantarum* LB1). Most genomes were assembled to the "Contig genome" level, with two complete genomes reported for *Pediococcus acidilactici* AF1 and *Enterococcus thailandicus* TH02. The count of contigs per genome displayed considerable variation, spanning from a minimal count of one contig for complete genomes and some contig genomes, to as many as 109 contigs observed in *L. plantarum* LB1. The longest contig reached 3,167,292 bp in *Lactiplantibacillus pentosus* CLP10, while the shortest among the longer contigs measured 646,627 bp in *L. plantarum* S1, indicating notable variability in maximum contig lengths. The N50 values, which indicate the length of the shortest contig at 50% of the total genome length, ranged from 326,094 bp in *L. plantarum* S1 to 3,226,988 bp in *L. plantarum* UK35. The L50 contig count, representing the smallest number of contigs that make up 50% of the genome, was as low as 1 for several samples, indicating highly contiguous assemblies (Supplementary Table 2). The assembly quality of each strain was evaluated using FastQC and reported in Supplementary Table 2.

Safety assessment

All strains were screened against the CARD and VFDB database to evaluate their safety. The findings revealed that none of the isolated strains harbored harmful genes linked to significant health risks, confirming their suitability for potential probiotic uses. In *L. plantarum* LB2, the arsD gene, encoding the arsenite efflux transporter metallo-chaperone ArsD, was identified with a 54.29% identity to the reference sequence, suggesting potential resistance to arsenic stress. In *Weissella paramesenteroides* WP12, the abc-f gene, encoding the ABC-F type ribosomal protection protein and conferring resistance to macrolides, was detected with a 62.85% identity. For *Enterococcus thailandicus* TH02, the *abc-f* gene was found with identities of 35.21% and 54.21%, indicating possible multiple occurrences of this resistance gene within the strain. Additionally, the *aac(6')-I* gene, encoding the aminoglycoside 6'-N-acetyltransferase associated with resistance to aminoglycosides, was identified with a 73.74% identity (Figure 1).

This figure illustrates the presence of AMR (Antimicrobial Resistance) and STRESS genes identified in various probiotic organisms. The chart shows the percentage identity of specific genes associated with antimicrobial resistance (*aac(6')-I* and *abc-f*) and stress response (*arsD*) in different probiotic strains.

Genetic characterization and Probiotic potential of bacterial strains

A comprehensive genetic analysis of various bacterial genomes reveals a large array of probiotic marker genes, highlighting their potential to enhance host health. Key strains such as *L. plantarum* S1, *P. acidilactici* CLF11, and

L. pentosus CLP10 exhibit genes for acid stress resistance, adhesion, antioxidant activity, bile resistance, carbohydrate metabolism, folate synthesis, immunomodulation, and vitamin synthesis.

Other notable strains including *L. plantarum* LB2 and LC3, *P. acidilactici* PA14, *P. stilesii* PA02, *L. fermentum* LF9, *W. paramesenteroides* WP12, *and E. thailandicus* TH02, feature a similar genetic profile, encompassing stress resistance, adhesion mechanisms, antioxidant properties, bile resistance, enzyme activity, folate synthesis, and immunomodulatory functions.

This genetic characterization points out the diverse probiotic functionalities of these strains, highlighting their potential to support host health through stress management, metabolic support, and immune modulation.

Discussion

The identification of potential probiotic strains enables researchers the great potential for further research and use. The screening of a large number of probiotics simultaneously provides advantages in various fields of study, whether gastrointestinal health, immune modulation, and metabolic support[7,21,22]. All identified strains in this study could be developed into probiotic products or supplemented into functional foods, which offer natural, safe, and effective solutions to promote human and animal well-being. Moreover, the comprehensive genomic characterization of probiotics provides solid support for further studies. It is probably enabling accurate formulation and targeted probiotic therapy. For example, the ability to tolerate acid tolerance and bile in different conditions indicates that these strains could survive and function effectively in the gastrointestinal tract of living organisms.

In this study, we obtained varying genome sizes across the different strains, ranging from 1,845,929 bp to 7,139,934 bp. This indicates the diversity of probiotic genomes and the potential differences in functional capabilities among them. For the assembly level, most of the strains in this study are assembled to the "Contig genome" level, which might occur due to the limitations of assembly software or the incompleteness of the length of DNA obtained from the DNA extraction step. However, two strains, including *P. acidilactici* AF1 and *E. thailandicus* TH02, achieved complete genome assembly, which typically provides more comprehensive and accurate genetic information. Moreover, the complexity of the target genome and contamination directly impact the assembly process. For instance, we encountered a 7.1 Mbp genome comprising two genomes of distinct probiotics: *L. plantarum* and *Weissella confusa* A similar scenario was observed in an

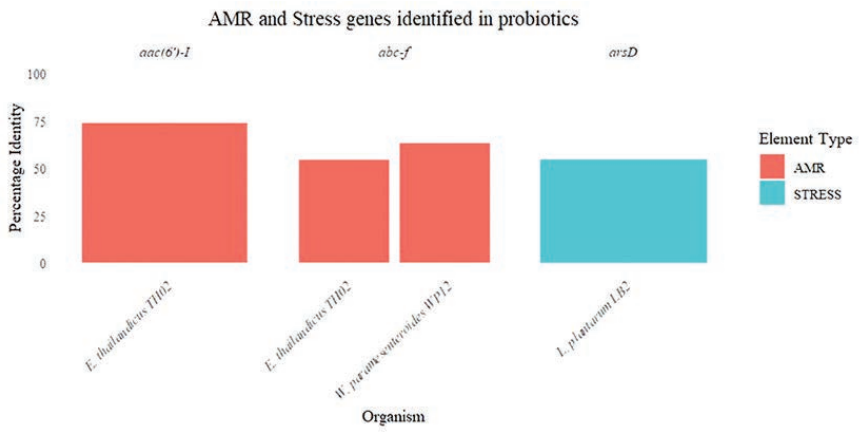

Figure 1 antimicrobial resistance characteristics in probiotic strains

unclassified sample, later identified as *P. pentosaceus* and *Weissella confusa* as well. Therefore, a higher number of contigs often suggests fragmentation in the assembly, which can complicate downstream analyses and interpretations. The assembly quality of each strain was evaluated using FastQC, as illustrated in Supplementary Table 2. Quality assessment is essential for identifying the quality and completeness of sequencing data, for example, low-quality reads that may impact the accuracy and reliability of the genome assemblies.

In order to identify and characterize appropriate probiotics to use for promoting health, several processes are required to determine their function[23]. In the past, probiotics need to be tested for phenotypes of antimicrobial resistance, virulence factors, stability, and hydrolysis activities. All these steps are time-consuming and require labor and resources. Therefore, this study aimed to screen and evaluate the probiotic properties of strains isolated from fermented foods and animal feces using nanopore sequencing and bioinformatic analysis. The results show that nanopore sequencing is an efficient tool for probiotic identification and characterization. The insights and genomic profiles obtained from the high throughput and accurate technique confirmed the presence of probiotic genes as well as ensured the safety of the candidate strains. Furthermore, nanopore sequencing provides several advantages to overcome traditional methods in terms of probiotic identification and characterization. Traditional techniques, such as culturing with selective media, biochemical tests, and PCR, often involve many processes which might lead to misidentification[24]. These methods typically focus on phenotypic characteristics and limited genetic markers, which can restrict the depth and accuracy of probiotic screening. In contrast, nanopore sequencing offers a high-throughput, real-time sequencing approach that provides long-read sequences suitable for whole-genome analysis[25]. In addition, ONT reveals comprehensive genomic information which enables a thorough assessment of each strain's potential benefits and safety. Insight in genomic data ensures that all genetic markers are evaluated, thus providing a more accurate view of the probiotic potential of the strains.

The 23 probiotic strains exhibit a range of beneficial properties that have significant clinical implications (Figure 2). These include enhancing gastrointestinal health by promoting a balanced gut microbiota, improving digestion, and preventing infections[26]. They also support immune system modulation, which can help in reducing the severity of infections and inflammatory responses. Additionally, the strains contribute to metabolic health by aiding in nutrient absorption and synthesizing essential vitamins[23]. Their antioxidant activities can protect against cellular damage, and their adhesion capabilities ensure they can effectively colonize the gut[27]. Overall, these strains hold promise for improving overall health and preventing various diseases, making them valuable for developing functional foods and therapeutic probiotics.

Oxford Nanopore Sequencing Technologies, particularly the MinION platform, offers fast and accurate identification of probiotic strains, making it a promising tool for future applications[28]. Its ability to generate long reads provides comprehensive genomic analysis crucial for characterizing these strains. However, considerations such as the ongoing costs of reagents and disposable equipment, as well as the need for specialized expertise to operate the technology and analyze data[29], must be addressed. While ONT is cost-effective in terms of initial investment and long-term use, ensuring financial feasibility and access to trained personnel are essential for its broader adoption in probiotic screening.

In conclusion, while ONT involves higher costs and requires technical expertise, its comprehensive, rapid, and scalable nature makes it suitable for the identification and screening of probiotic bacteria, especially in research settings where detailed genomic insights are essential.

Figure 2 List of probiotic marker genes identified in the genome presents the results of probiotic gene analysis in each sample, indicating their functional categories

Conclusion

Utilizing nanopore sequencing for probiotic screening not only offers a thorough assessment of probiotic properties but also introduces a rapid, precise, and scalable methodology. This approach overcomes the constraints of conventional techniques and facilitates numerous studies investigating probiotic health benefits. Integrating nanopore sequencing into probiotic research presents a notable advantage, bolstering capabilities for discovering and crafting effective probiotic interventions. Moreover, this technology serves as a cornerstone in developing next-generation probiotics that are safer, more efficacious, and assist in addressing specific health requirements.

Acknowledgement

This research was supported by Graduate scholarship, Faculty of Medicine, Prince of Songkla University, whose financial assistance made this research possible. We are also grateful to Prince of Songkla University for providing the infrastructure and resources needed to conduct this study.

Conflict of interest

There are no potential conflicts of interest to declare.

References

1. Sood A, Midha V, Makharia GK, Ahuja V, Singal D, Goswami P, Tandon RK. The probiotic preparation, VSL#3 induces remission in patients with mild-to-moderately active ulcerative colitis. Clin Gastroenterol Hepatol 2009;7:1202-9.
2. Niv E, Naftali T, Hallak R, Vaisman N. The efficacy of *Lactobacillus reuteri* ATCC 55730 in the treatment of patients with irritable bowel syndrome—a double blind, placebo-controlled, randomized study. Clinical Nutrition 2005;24:925-31.
3. Kadooka Y, Sato M, Imaizumi K, Ogawa A, Ikuyama K, Akai Y, et al. Regulation of abdominal adiposity by probiotics (*Lactobacillus gasseri* SBT2055) in adults with obese tendencies in a randomized controlled trial. Eur J Clin Nutr 2010;64:636-43.
4. Kadooka Y, Sato M, Ogawa A, Miyoshi M, Uenishi H, Ogawa H, et al. Effect of *Lactobacillus gasseri* SBT2055 in fermented milk on abdominal adiposity in adults in a randomised controlled trial. Br J Nutr 2013;110:1696-703.
5. Chao L, Liu C, Sutthawongwadee S, Li Y, Lv W, Chen W, et al. Effects of probiotics on depressive or anxiety variables in healthy participants under stress conditions or with a depressive or anxiety diagnosis: a meta-analysis of randomized controlled trials. Front Neurol 2020;11:421.
6. Kober MM, Bowe WP. The effect of probiotics on immune regulation, acne, and photoaging. Int J Womens Dermatol 2015;1:85-9.
7. Teng Y, Huang Y, Danfeng X, Tao X, Fan Y. The Role of Probiotics in Skin Photoaging and Related Mechanisms: A Review. Clin Cosmet Investig Dermatol 2022;15:2455-64.
8. Momin ES, Khan AA, Kashyap T, Pervaiz MA, Akram A, Mannan V, et al. The effects of probiotics on cholesterol levels in patients with metabolic syndrome: a systematic review. Cureus 2023;15:e37567.
9. Spanu T, Sanguinetti M, Ciccaglione D, D'Inzeo T, Romano L, Leone F, Fadda G. Use of the VITEK 2 system for rapid identification of clinical isolates of Staphylococci from bloodstream infections. J Clin Microbiol 2003;41:4259-63.
10. Ashfaq MY, Da'na DA, Al-Ghouti MA. Application of MALDI-TOF MS for identification of environmental bacteria: A review. J Environ Manage 2022;305:114359.
11. Jarvinen AK, Laakso S, Piiparinen P, Aittakorpi A, Lindfors M, Huopaniemi L, et al. Rapid identification of bacterial pathogens using a PCR- and microarray-based assay. BMC Microbiol 2009;9:161.
12. Tyler AD, Mataseje L, Urfano CJ, Schmidt L, Antonation KS, Mulvey MR, Corbett CR. Evaluation of oxford nanopore's minion sequencing device for microbial whole genome sequencing applications. Sci Rep 2018;8:10931.
13. Megur A, Daliri EB, Balnionyte T, Stankeviciute J, Lastauskiene E, Burokas A. In vitro screening and characterization of lactic acid bacteria from Lithuanian fermented food with potential probiotic properties. Front Microbiol 2023;14:1213370.
14. ZymoBIOMICS™ DNA Miniprep Kit [homepage on the Internet]. California: Zymo Research Europe; 1994 [cited 2024 Jun 14]. Available from: https://zymoresearch.eu/products/zymobiomics-dna-miniprep-kit.
15. Rapid sequencing DNA V14 - barcoding (SQK-RBK114.24 or SQK-RBK114.96) [homepage on the Internet]. Oxford: Oxford

Nanopore Technologies; 2008 [cited 2024 Jun 14]. Available from: https://nanoporetech.com/

16. Petit RA, 3rd, Read TD. Bactopia: a flexible pipeline for complete analysis of bacterial genomes. mSystems 2020;5.
17. Brown J, Pirrung M, McCue LA. FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. Bioinformatics 2017;33:3137-9.
18. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29:1072-5.
19. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya Amogelang R, Wlodarski MA, et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the comprehensive antibiotic resistance database. Nucleic Acids Res 2023;51:D690-9.
20. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, Jin Q. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res 2005;33:D325-8.
21. Cristofori F, Dargenio VN, Dargenio C, Miniello VL, Barone M, Francavilla R. Anti-inflammatory and immunomodulatory effects of probiotics in gut inflammation: a door to the body. Front Immunol 2021;12:578386.
22. Hungin APS, Mitchell CR, Whorwell P, Mulligan C, Cole O, Agreus L, et al. Systematic review: probiotics in the management of lower gastrointestinal symptoms-an updated evidence-based international consensus. Aliment Pharmacol Ther 2018;47:1054-70.
23. Latif A, Shehzad A, Niazi S, Zahid A, Ashraf W, Iqbal MW, et al. Probiotics: mechanism of action, health benefits and their application in food industries. Frontiers in Microbiology 2023;14.
24. Suwannaphan S. Isolation, identification and potential probiotic characterization of lactic acid bacteria from Thai traditional fermented food. AIMS Microbiol 2021;7:431-46.
25. Ciuffreda L, Rodriguez-Perez H, Flores C. Nanopore sequencing and its application to the study of microbial communities. Comput Struct Biotechnol J 2021;19:1497-511.
26. Verna EC, Lucak S. Use of probiotics in gastrointestinal disorders: what to recommend? Therap Adv Gastroenterol 2010;3:307-19.
27. Han S, Lu Y, Xie J, Fei Y, Zheng G, Wang Z, et al. Probiotic gastrointestinal transit and colonization after oral administration: a long journey. Front Cell Infect Microbiol 2021;11:609722.
28. Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol 2016;17:239.
29. Zheng P, Zhou C, Ding Y, Liu B, Lu L, Zhu F, Duan S. Nanopore sequencing technology and its applications. MedComm (2020). 2023;4:e316.

Supplementary Table 1 isolation and identification of microorganisms from fermented foods and animal feces using MALDI-TOF MS

Date	Isolation source	Sample ID	Location	MALDI TOF
26-Apr-23	fermented food (Khai-Khrop-Songkhla)	PK3	Songkhla, Thailand	*Pichia kudriavzevii*
26-Apr-23	fermented food (Hoi-Dong)	LB2	Songkhla, Thailand	*Lactobacillus brevis**
26-Apr-23	fermented food (Phak-Kat-Dong)	PK2	Songkhla, Thailand	*Pichia kudriavzevii*
26-Apr-23	fermented food (Phak-Sian-Dong)	LB1	Songkhla, Thailand	*Lactobacillus brevis**
11-May-23	fermented food (Pu-Dong-Khem)	PK1	Songkhla, Thailand	*Pichia kudriavzevii*
11-May-23	fermented food (Kha-nom-chin)	LF5	Songkhla, Thailand	*Limosilactobacillus fermentum**
11-May-23	fermented food (Kha-nom-chin)	UK21	Songkhla, Thailand	No Organism Identification Possible*
11-May-23	fermented food (Kung-Som)	TH02	Songkhla, Thailand	*Enterococcus faecium**
19-May-23	fermented food (Pla-Peng-Deng)	CLF11	Songkhla, Thailand	*Companilactobacillus farciminis**
19-May-23	fermented food(Sato-Dong)	CLP10	Songkhla, Thailand	*Companilactobacillus farciminis**
19-May-23	fermented food (Pla-Peng-Deng)	PA14	Songkhla, Thailand	*Pediococcus acidilactici**
19-May-23	fermented food (Kung-Som)	DR1	Songkhla, Thailand	*Diutina rugosa*
19-May-23	fermented food (Namtan-Tanot)	ZB1	Songkhla, Thailand	*Zygosaccharomyces bailii*
19-May-23	fermented food (Kung-Som)	LP8	Songkhla, Thailand	*Lactiplantibacillus plantarum**
19-May-23	fermented food(Nam-Bu-Du)	WP12	Songkhla, Thailand	*weissella paramesenteroides**
19-May-23	fermented food (Peng-Khao-Mak)	UK35	Songkhla, Thailand	No Organism Identification Possible*
19-May-23	fermented food (Phutsa-Dong)	PA02	Songkhla, Thailand	No Organism Identification Possible*
19-May-23	fermented food (Plara)	AF1	Songkhla, Thailand	*Pediococcus acidilactici**
19-May-23	animal feces (Goat)	PD02	Songkhla, Thailand	*Pediococcus pentosaceus**
19-May-23	fermented food (Nomai-Dong)	NM1	Songkhla, Thailand	*Limosilactobacillus fermentum**
19-May-23	fermented food (Peng-Mak-Khanomchin)	LF9	Songkhla, Thailand	*Limosilactobacillus fermentum**
6-Oct-23	animal feces (Pig)	AP1	Songkhla, Thailand	*Escherichia coli*
6-Oct-23	animal feces (Pig)	BBP1	Songkhla, Thailand	*Escherichia coli*
6-Oct-23	fermented food (Peng-Khao-Mak)	P1	Songkhla, Thailand	*Pediococcus pentosaceus**
6-Oct-23	fermented food (Makham-Dong)	E4	Songkhla, Thailand	*Pediococcus pentosaceus**
6-Oct-23	fermented food (TonHom-Dong)	H2	Songkhla, Thailand	*Pediococcus pentosaceus**
6-Oct-23	fermented food (Peng-Khao-Mak)	LC3	Songkhla, Thailand	*Lactobacillus curvatus**
6-Oct-23	fermented food (Plara)	PK3	Songkhla, Thailand	*Pichia kudriavzevii*
6-Oct-23	fermented food (Peng-Khao-Mak)	M1	Songkhla, Thailand	*Pediococcus pentosaceus**
6-Oct-23	fermented food (Peng-Khao-Mak)	WC2	Songkhla, Thailand	*Weissella cibaria**
6-Oct-23	fermented food (Mamuang-Che-Im)	S1	Songkhla, Thailand	*Lactiplantibacillus plantarum**

*The 23 non-pathogenic strains with asterisk were selected for whole-genome sequencing (WGS), MALDI-TOF MS=matrix-assisted laser desorption/ionization time-of-flight mass spectrometer

Supplementary Table 2 The predicted amount of data and the obtained data using nanopore sequencing

Sample ID	MALDI TOF			Nanopore sequencing		
	Organism	Reference genome size (Mdp)	Predicted data (Mdp)	Identification	Genomesize (Mdp)	Total bases
S1	*Lactiplantibacillus plantarum*	3.2–3.4	128–136	*Lactiplantibacillus plantarum*	3.33	32,549,758.00
CLP10	*Companilactobacillus farciminis*	2.6	104	*Lactiplantibacillus pentosus*	3.96	64,999,056.00
PA14	*Pediococcus acidilactici*	1.7–2.1	68–84	*Pediococcus acidilactici*	2.00	257,132,757.00
CLF11	*Companilactobacillus farciminis*	2.6	104	*Pediococcus acidilactici*	1.85	258,673,021.00
LB2	*Lactobacillus brevis*	2.2–2.7	88–108	*Lactiplantibacillus plantarum*	5.92	215,356,041.00
LP8	*Lactiplantibacillus plantarum*	3.2–3.4	128–136	*Lactiplantibacillus plantarum*	3.34	124,227,421.00
LF5	*Limosilactobacillus fermentum*	1.9–2.1	76–84	*Limosilactobacillus fermentum*	5.27	174,855,475.00
WP12	*weissella* paramesenteroides	2	80	*Weissella paramesenteroides*	2.07	159,494,346.00
UK35	*No Organism Identification Possible*	–	NA	*Lactiplantibacillus*	5.61	387,682,654.00
E4	*Pediococcus pentosaceus*	1.7–1.8	68–72	*Lactiplantibacillus pentosus*	4.28	127,388,818.00
H2	*Pediococcus pentosaceus*	1.7–1.8	68–72	*Pediococcus pentosaceus*	4.32	227,212,806.00
LB1	*Lactobacillus brevis*	2.2–2.7	88–108	*Lactiplantibacillus plantarum*	7.13	127,166,737.00
LC3	*Lactobacillus curvatus*	1.8–2.1	72–84	*Lactiplantibacillus plantarum*	3.44	81,159,467.00
PA02	No Organism Identification Possible	–	NA	*Pediococcus stilesii*	2.14	317,321,833.00
AF1	*Pediococcus acidilactici*	1.7–2.1	68–84	*Pediococcus acidilactici*	1.9	129,667,473.00
PD02	*Pediococcus pentosaceus*	1.7–1.8	68–72	*Pediococcus pentosaceus*	1.86	336,941,114.00
NM	*Limosilactobacillus fermentum*	1.9–2.1	76–84	*Limosilactobacillus fermentum*	2.48	170,893,552.00
M1	*Pediococcus pentosaceus*	1.7–1.8	68–72	Unclassified Bacteria	4.34	290,784,164.00
WC2	*Weissella cibaria*	2.5–2.6	100–104	*Weissella confusa*	2.52	276,216,732.00
P1	*Pediococcus pentosaceus*	1.7–1.8	68–72	*Pediococcus pentosaceus*	1.84	198,661,939.00
LF9	*Limosilactobacillus fermentum*	1.9–2.1	76–84	*Limosilactobacillus fermentum*	2.25	94,984,506.00
UK21	*No Organism Identification Possible*	–	NA	*Lactiplantibacillus plantarum*	3.88	129,335,023.00
TH02	*Enterococcus faecium*	2.5–2.7	100–108	*Enterococcus thailandicus*	2.77	228,197,849.00

MALDI TOF=matrix-assisted laser desorption/ionization time-of-flight