Pakistan J. Zool., vol. 56(4), pp 1827-1835, 2024 **DOI**: https://dx.doi.org/10.17582/journal.pjz/20210708190732

Genome Characterization of Probiotic *Lactobacillus delbrueckii* Subsp. *bulgaricus* Strain NDO2 Isolated from Traditional Yogurt using High Throughput Next Generation Sequencing

Mati Ullah[1]*, Muhammad Rizwan[2], Ali Raza[3], Xianlin Zhao[4], Yanling Sun[4], Sarah Gul[5], Muhammad Ihtesham Waheed[6], Muhammad Nadeem Khan[7], Alvina Gul[8], Sami Ullah Jan[9]* and Chao Huang[4]*

[1]*Department of Microbiology, Faculty of Veterinary & Animal Sciences, The University of Agriculture, Dera Ismail Khan, Pakistan*
[2]*College of Animal Sciences and Technology, Anhui Agricultural University, Hefei, Anhui, China*
[3]*Hefei National Laboratory for Physical Sciences at Microscale, School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, China*
[4]*Biotechnology Research Center, Shandong Academy of Agricultural Sciences, Jinan, China*
[5]*Department of Biological Sciences, Faculty of Basic and Applied Sciences, International Islamic University, Islamabad, Pakistan*
[6]*Department of Plant Breeding and Genetics, University of Agriculture, Peshawar*
[7]*Department of Microbiology, Faculty of Biological Sciences, Quaid-I-Azam University, Islamabad, Pakistan*
[8]*Atta-Ur-Rahman School of Applied Biosciences, National University of Sciences and Technology, H-12 Campus, Islamabad, Pakistan*
[9]*Department of Bioinformatics and Biosciences, Faculty of Health & Life Sciences, Capital University of Science and Technology, Islamabad, Pakistan*

Article Information
Received 08 July 2021
Revised 20 January 2022
Accepted 04 February 2022
Available online 07 April 2023
(early access)
Published 11 June 2024

Authors' Contribution
CH and MU conceived the study and designed the project. MR and MIW collected the samples. MU conducted the experiments. MU and CH carried out the bioinformatics analysis. MU, MR, AR, XZ, YS, SG, SUJ and MIW analyzed the data. MU and MNK wrote and prepared the manuscript. AG and CH revised the manuscript. All authors read and approved the final manuscript.

Key words
***Lactobacillus delbrueckii Bulgaricus*, Strain NDO2, Next-generation sequencing, Yogurt**

ABSTRACT

Probiotic bacteria present in fermented milk and its products have a long history of safety and health associated uses. If bacteria isolated from yogurt is characterized technologically and gnomically, it could be used in fermentation for selective purposes. In the current study, we have comprehensively analyzed the genome of *Lactobacillus delbrueckii* subsp. *bulgaricus NDO2* isolated from traditional yogurt samples using the next-generation sequencing approach. The GC contents of the selected genome were 45.3%. Among the total 2284 genes identified, 970 genes similar in three genomes make the core genome portion. The subsystems, category distribution analysis revealed "carbohydrates" was the largest subsystem, followed by "protein metabolism". The eggNOG functional analysis revealed that the COGs associated with "Carbohydrate transport and metabolism were dominant followed by those associated with "Replication, recombination, and repair" while the biological pathway analysis by the Kyoto encyclopedia of genes and genomes (KEGG) database identified the metabolic and biosynthesis of secondary metabolites pathway as the dominant one. These findings provide a broader genomic understanding of the famous probiotics bacterial species for general as well as commercial applications. This study could also be used basis for similar studies.

* Corresponding author: samiullahjan@gmail.com, 67731020@qq.com; matiqau@gmail.com
0030-9923/2024/0004-1827 $ 9.00/0

INTRODUCTION

Yogurt, the fermented milk product is a famed dairy item around the globe and is usually consumed in various forms among people of all ages (Fiszman and Salvador, 1999).This fermented food and its related products contain viable lactic acid bacteria (LAB) which are linked with numerous health benefits (Staffolo *et al*., 2004). The beneficial effect of yogurts associated with the

well-being of humans has a long history. Several scientific investigations have documented the positive impact of yogurt cultures on gut metabolism (Meydani and Ha, 2000; Adolfsson *et al*., 2004). Moreover, the bacteria in yogurts can improve lactose digestion (Lerebours *et al*., 1989; Labayen *et al*., 2001), affect the intestinal transit time (Conway *et al*., 1987), and improve the immune system of the gut (Van de Water *et al*., 1999; Aldinucci *et al*., 2002). Lactic acid bacteria (LAB) are the main and prominent microbial component of dairy-related fermented foods. The member of this important group of bacteria is used as a starter culture in the fermentation process. These are gram positive bacteria and are present as indigenous microflora in milk and yogurts. This specific group of bacteria contains numerous species from the genus *Lactobacillus, Leuconostoc, and Lactococcus* (Tserovska *et al*., 2002; Papadimitriou *et al*., 2015). LAB serves as a natural source of preservation, associated with the texture, flavor, and nutritional value of the fermented foods (Atrih *et al*., 2001; Messens and De, 2002; Mohd Adnan and Tan, 2007).

Probiotics are live microorganisms which when administered in adequate amounts confer a health benefit to the host (Hill *et al*., 2014; Papadimitriou *et al*., 2015). Numerous strains of different probiotic bacterial species are associated with human health. The reduction of cholesterol level, modulation of immunity, and effective treatment of atopic dermatitis, diarrhea, and Crohn's disease are linked with probiotics (Reid, 1999).

Lactobacillus delbrueckii subsp. *Bulgaricus* is a gram positive, facultatively anaerobic, and industrially important lactic acid bacteria. This species is not native to humans, however, can survive the gastric transit (Mater *et al*., 2005). It has been used for decades in combination with *Streptococcus thermophilus* and has been very important industrially ever since. Both species have synergistic association, but thickness and aroma of the yogurt is linked with *Lactobacillus delbreuckii*'s pH lowering and acetaldehyde producing abilities (Gezginc *et al*., 2015). Particularly, *L. delbreuckii* subsp *bulgaricus* NDO2, due to its moderate acidity, high viscosity and excellent water holding capacity (Sun *et al*., 2011). In combination with other gastrointestinal flora including, but not limited to *Lactobacillus casei*, *Lactobacillus acidophilus*, and *Lactobacillus rhamnosus*, *L. delbreruki* has been used to improve nutritional values of the yogurt (Heller, 2001; Vaughan *et al*., 2002). Some strains of *L. delreuckii* has been reported as Bacteriocin producers, which suggests that strains of this specie have defensive properties and can be used as preservatives and therapeutic agents (Boris *et al*., 2001; Simova *et al*., 2008). The *L. delbreuckii* subspp *bulgricus* NDO2 has been known resistant to freeze drying temperature and desiccation which further confirms it a best starter candidate for industrial use (Shao *et al*., 2014). Initial results of *L. delbreucki* research in recent years claim that this specie has potential to manage many medical complications like fatty liver disease, flue, Antibiotic associated diarrhoea, Inflammatory bowel disease, eczema, hay fever, colic, tooth decay and enterocolitis (Hempel *et al*., 2012; Ghouri *et al*., 2014; Ouwehand *et al*., 2016).

No matter how many benefits a bacterium may have, according to FDA, it cannot be introduced to the food chain for use until its genome is known (Degnan, 2008). Genomic studies of bacteria are vital as it authenticate the experimental work done in laboratories, enlightens the possibilities for future research and discovers safety fears. Bacteria are genetically flexible and have high adoptability with respect to habitual and nutritional situations. Hence, making their characteristic strain specific. Because, Lactic acid bacteria's genome is fragile and is prone to shredding, it is possible that its genome may have time dependent variation. The application of high throughput next-generation sequencing in microbial genome characterization has gained much attention over the last few years. In the current study, we used this emerging sequencing approach to comprehensively analyze the genome of industrially and medically important probiotic *L. delreuckii* NDO2 bacteria isolated from traditional yogurt sold in the Karak district of Khyberpakhtoonkhwa, Pakistan. The findings of the current study can provide a detailed insight into the genome of *L. delbreuckii* NDO2 as well as advance the dairy research in terms of starter culture characterization using the next-generation sequencing platform.

MATERIALS AND METHODS

Yogurt samples

Samples of cow's milk artisanal yogurt were collected from local vendors in the Karak district of Khyber Pakhtunkhwa, Pakistan. Sterile sampling bottles with and ice pack were used for collection of the samples. Soon after collection the samples were sent to the laboratory.

Isolation and identification of probiotic bacterial species

The yogurt samples were aseptically homogenized, and one milliliter of the sample was diluted in 9 milliliters of phosphate-buffered saline (PBS) followed by serial dilutions. Different dilutions were plated on Man–Rogosa–Sharpe (MRS) pH 5.2 agar plates. Plates were incubated in triplicate under anaerobic conditions at 37, and 45 °C for 24–72 h. The Yuejin anaerobic incubator was used for probiotic bacterial growth (Shanghai Yuejin Medical Instruments Co., Ltd, Shanghai, China). Colonies from plates were sub-cultured till obtaining pure isolated

colonies. The pure culture isolates were used for Gram reaction, cell morphology, etc.

DNA extraction

The pure bacterial colony from the agar plate was grown overnight on MRS broth followed by DNA extraction using a Tiangen bacterial DNA isolation kit (Tiangen, Beijing, China). Further, Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA, USA) was used to quantify the extracted DNA.

DNA library preparation and sequencing

DNA Libraries for high-throughput sequencing were made by commercially available Vazyme TruePrep™ DNA Library Prep Kit V2 for Illumina (Vazyme Biotech, Nanjing, China). The sequencing of the sample libraries was achieved on an Illumina HiSeq-2000 sequencing platform (BGI-Shenzhen, Shenzhen, China).

Bioinformatics analysis

Trimming, quality control, and de novo assembly

Trimming of the sequencing reads was completed by means of the Trimmomatic version 0.38 (Bolger *et al*., 2014). The contaminated reads were screened through BWA aligned to the reference sequence. The assembly of the high-quality sequencing reads was made with shovill (version 1.1.0, https://github.com/tseemann/shovill) assembler. The de novo assembly results were evaluated using the Quast software (Gurevich *et al*., 2013). DNA Plotter (Carver *et al*., 2009), an interactive Java application was used for generation of circular DNA plots. The draft genome of isolate LB2 was taken as the reference genome, and randomly connected the Contig sequences of the LB2 genome to form a pseudo-single genome sequence. The Prokka program was used for annotations and Align the Y14 and Y16 genomes to the reference genomic group constructed by LB2, using a minimum similarity of 70%, and convert the comparison results into GFF format. The results were imported in the form of image.

Genome annotation

The assembled genome was annotated through the Prokka (version 1.14.5, 10.1093/bioinformatics/btu153) pipeline. The subsystem identification in the genome and functions assigning was carried out through the protein-coding sequences (CDS). The comparison of the assembled genome sequences was assessed in the SEED Subsystem (Overbeek *et al*., 2014). The evaluation of orthologous genes was performed via clusters of orthologous genes (COGs) (Galperin *et al*., 2015) and eggnog (Huerta-Cepas *et al*., 2016). To examine the high level functions of the three isolates, the Kyoto encyclopedia of genes and genomes (KEGG) database (https://www.genome.jp/kegg/) was retrieved.

Genome analysis

The pan-genome analysis of the three genomes was done by roary (version 3.13.0, 10.1093/molbev/msz284) pipeline. Briefly, the protein sequences from gff 3 files were extracted to a reduced set of protein sequences followed by an all against all comparison with BLASTP. The sequence identity was set 98% with e-value 1E-6. Based on BLASTP similarity score, sequences were clustered with a Markov cluster (MCL) algorithm. Further, a concatenated core gene alignment sequence was produced using PRANK (10.1007/978-1-62703-646-7_10).

The raw data used in the current study have been deposited in NCBI and allocated the accession number of PRJNA664263.

RESULTS

Sequencing data and assembled genome properties

The summary and statistics of the sequence data and assembly results is shown in Table I. The number of clean reads used in the analysis was in the range of 2303582 to 6604528, with a mean value of 4848434.29. The GC contents were in the range between 38.79 to 49.09%, with 45.3% is the mean value. The coverage and GC contents of each identified species/strains in the corresponding sample are shown in Figure 1 as well as in Figure 2. The assembled contigs from the genome sequencing data of the three bacterial genomes vary significantly and were in the range of 116 (lowest) for sample NY4 and 1605 (highest) for sample LB2. The other important genome features characterized in the current analysis are summarized in Table I.

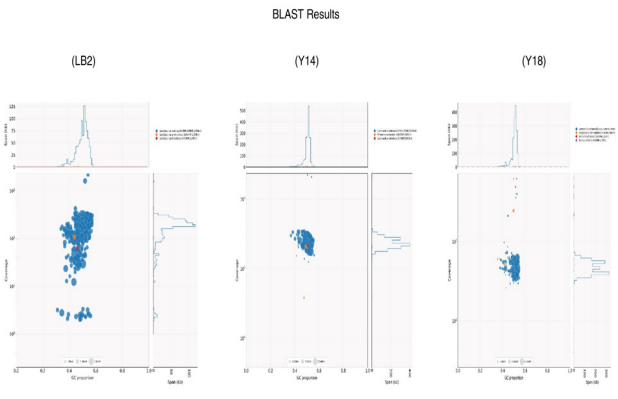

Fig. 1. Blob plots indicate the results of BLAST matches of each of the three assembled genomes. The colors indicate BLAST matches to different species of bacteria. The proportion of GC contents in each assembled genome is also indicated.

Table I. Summary of sequencing data and assembled genome properties.

Sample No	1	2	3
Library name	LB2	Y14	Y18
Species/strain	*Lactobacillus delbrueckii* subsp. bulgaricus ND02	*Lactobacillus delbrueckii* subsp. bulgaricus ND02	*Lactobacillus delbrueckii* subsp. bulgaricus ND02
Raw Reads	6559393	6871279	5924954
Clean Reads	5668859	6604528	5717889
Raw base (G)	1.97	2.06	1.78
Clean Base (G)	1.7	1.98	1.72
GC Contents%	49.09	46.05	38.79
Contigs	1605	169	206
Bases	1643781	1687821	1719327
CDS	1580	1628	1653
Gene	1657	1755	1837
mRNA	1657	1755	1837
Misc_RNA	17	22	21
rRNA	3	7	14
Repeat region	3	1	–
tRNA	57	97	148
tmRNA	–	1	1

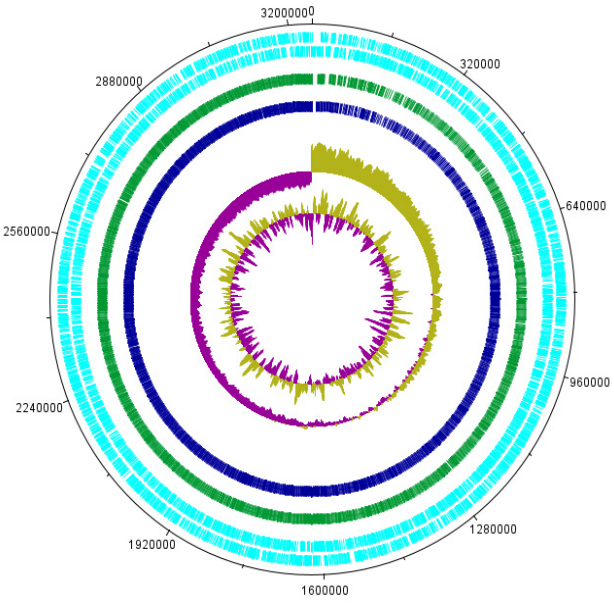

Fig. 2. Circular plot of *Lactobacillus delbrueckii* subsp. *Bulgaricus strain* NDO2 genome generated by DNA plotter. Circles indicate, from inside outwards: GC skew; GC content; Y18 Comparison results; Y14 comparison result; CDS sequence are (reverse frame); CDS sequence area (plus frame).

Comparative genome analysis

The annotated genomes of the three isolates of the *L. delbrueckii* subsp. *bulgaricus* strain NDO2 were mutually compared. The pan-genome analysis of the three isolates of *L. delbrueckii* subsp. *bulgaricus* NDO2 identified a total of 2284 genes among which 970 represent core genes and 1314 represent shell genes respectively. The three isolates of the species *L. delbrueckii* subsp. *bulgaricus* NDO2 share a similar core genome, however, regarding the accessory genes pattern, the isolates of Y14 and Y18 showed more similarity compare to LB2. The pattern of the number of conserved genes shows a continuous decline compare to the number of total genes as the number of genomes embedded in the analysis. Similarly, the number of both unique and new genes declined as the number of genomes were embedded into the analysis. Graphical summarized illustration of the comparative gene analysis is presented in Figure 3. Detail of the pangenome analysis is given in Supplementary Table SI.

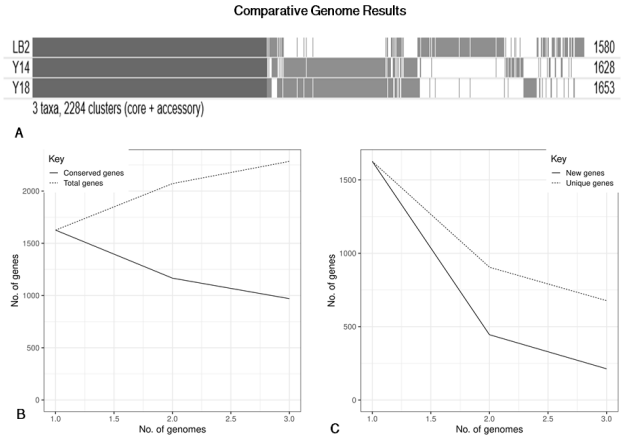

Fig. 3. Comparative genome analysis of the three *L. delbrueckii* subsp. *Bulgaricus* isolates. (A) Roary matrix. Pan-genome analysis of three *L. delbrueckii* subsp. *Bulgaricus* isolates annotated genomes. Roary produced the gene presence/absence matrix. A total of 2284 protein-coding gene sequence clusters produced by roary are shown in the figure. The presence of the gene is represented by grey color while the absence is shown by white color. The numbers represent clusters in each assembled annotated genome. (B) Diagram of conserved genes per number of genomes. (C) Diagram of new and unique genes per number of genomes.

Functional genome annotation

Subsystem category distribution

The subsystem category distribution of the dominant systems identified in the three annotated genomes using the SEED database. Among these subsystems, carbohydrates was the largest subsystem, followed by

protein metabolism. The other dominant systems were DNA metabolism and nucleoside and nucleotide (Fig. 4). Besides this, numerous other subsystems with different numbers were also identified. Their detail is provided in Supplementary Tables SII, SIII, SIV).

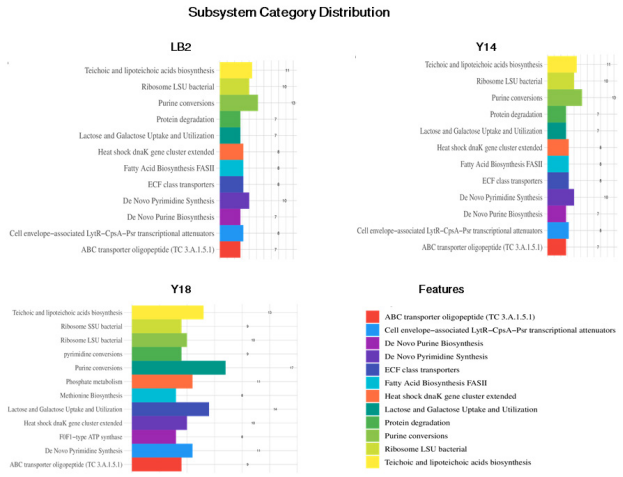

Fig. 4. The subsystem category distribution of the probiotic bacterial genome sequences according to the SEED database. The largest subsystem feature counts belong to carbohydrates, protein metabolism, DNA metabolism and nucleoside and nucleotide. The description of various subsystem features is represented in a different color below right.

Genome annotation with eggNOG database

The eggnog database, hosted by the European molecular biology laboratory, is a database of orthologous groups and function annotation of the genome. The functional study analysis of the three annotated genomes revealed that the COGs associated with carbohydrate transport and metabolism were dominant followed by those associated with Replication, recombination, and repair (Fig. 5). The other dominant categories with most COGs associated were amino acid transport and metabolism and translation ribosomal structure and biogenesis, respectively. For viewing the results in detail Supplementary data can be accessed (Supplementary Tables SV, SVI, SVII).

KEGG analysis

The information provided by the KEGG database is related to the high level functions for high throughput sequencing data regarding genetic information processing, metabolism, cellular processes, human diseases, environmental information processing, drug development, and organismal systems. Using the KEGG database pathways as a reference, the genome content in the current analysis were characterized to identify the biological pathways. Most gene contents were involved in metabolic pathways, biosyn thesis of secondary metabolites, microbial metabolism in diverse environments, and biosynthesis of antibiotics respectively (Fig. 6). More detail of the results is given in Supplementary Tables SVIII, SIX, SX.

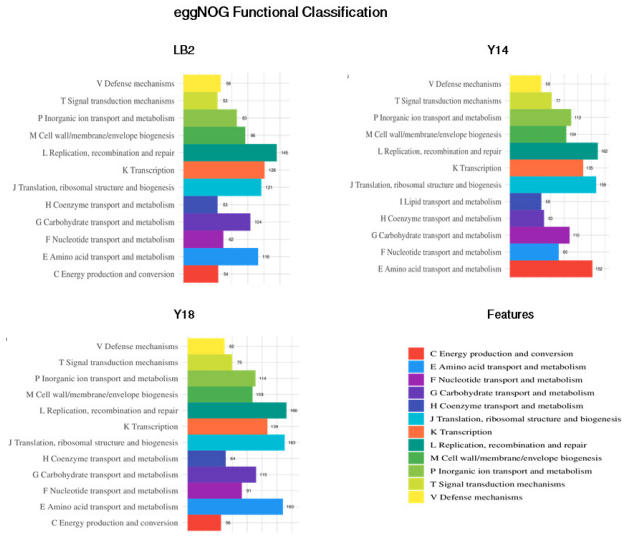

Fig. 5. Classification of the eggNOG annotation of the probiotic bacterial genome sequences using eggNOG. The numbers in the plot represent the genes assigned to eggNOG classification. The description of the various annotated eggNOG features is represented in various colors below right.

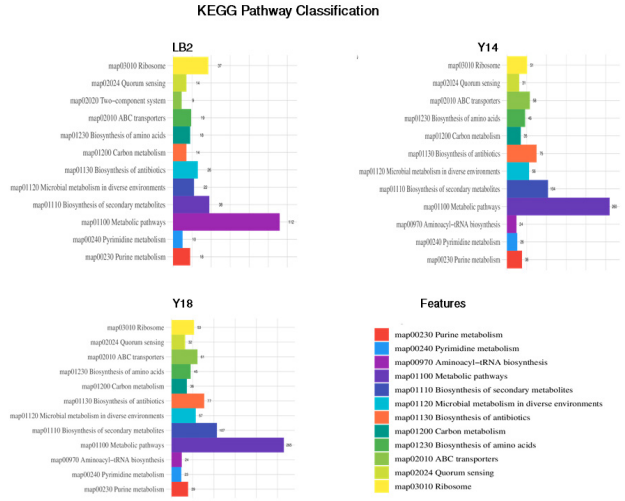

Fig. 6. Classification of the KEGG pathways based on the assembled genome of the three probiotic bacterial isolates. The different colors represent the various KEGG pathways identified in the current analysis.

DISCUSSION

Bacterial strains used as starter cultures for fermentation of yogurt are of inordinate curiosity because they have a direct effect on nutritional, therapeutic and commercial values of the final product. In dairy industry for fermentation of milk, commercially available strains of two bacteria, *S. thermophilus* and *L. bulgaricus* are most used. These strains can produce yogurt with necessary properties in a gainful way, but some wild strains to be discovered may be superior to them and upsurge the qualities of the product, reduce the expenses for yogurt manufacture, or yield new types of yogurt with different properties. To translate microbial potentials, complete genome sequencing of industrially important microorganisms holds vital significance. In order to understand and easily utilize important bacteria, whole genome sequencing of multiple species and re-sequencing of existing genomes are extremely needed.

In this study, we aimed to isolate *L. debreukii* subsp *bulgaricus* NDO2 strains from traditional yogurt and performed their whole genome sequencing for functional understanding and comparison. This study is also important because no other genome of this strain has yet submitted to the NCBI genome data base (Sun *et al*., 2011).

The number of clean reads used in the analysis was in the range of 2303582 to 6604528, with a mean value of 4848434.29. while in case of the initial reported strain NDO 2 total of reads were reported to be 8,778,388 that constituted complete genome of ND02 having 2,125,753 base pairs (Sun *et al*., 2011). The GC contents we analyzed for the genomes were in the range between 38.79 to 49.09%, with 45.3% the mean value. The published literature supports ours results as lactic acid bacteria have GC contents ranging from 32 to 52%/ (Mendes-Soares *et al*., 2014). While GC contents of the NDO2 which is used as a reference strain are 49.56% (Sun *et al*., 2011). The difference in percentage of GC content is directly linked with genomic size and structure, and ecological niche to which a microorganism belongs, thus it may vary even for strains (Šmarda *et al*., 2014). The assembled draft genome contains a mean of 1620 coding sequences that encodes 1749 genes. The submitted genome of early reported strain contains 2151 coding sequences with 2177 genes (Sun *et al*., 2011). The variation in coding sequences and gene number is also strain dependent trait.

The assembled contigs from the genome sequencing data of the three bacterial genomes vary significantly and were in the range of 116 (lowest) for sample NY4 and 1605 (highest) for sample LB2. The effect of contig number on the functional annotation has not been proven by the literature or there is any standard for the number of contigs. The variation in contig number is associated with the technology being used for sequencing and methodology used for assembly (Smits, 2019).

The pan-genome analysis of the three isolates of *L. delbrueckii* subsp. *bulgaricus* identified a total of 2284 genes among which 970 represent core genes and 1314 represent shell genes respectively (need to know functions of core genes and shell genes). The three isolates of the species *L. delbrueckii* subsp. *bulgaricus* share a similar core genome, however, regarding the accessory genes pattern, the isolates of Y14 and Y18 showed more similarity compare to LB2. The pattern of the number of conserved genes shows a continuous decline compare to the number of total genes as the number of genomes embedded in the analysis. Similarly, the number of both unique and new genes declined as the number of genomes were embedded into the analysis. Differences in the accessory genes is associated with strain specific characteristics (Pintara *et al*., 2020).

The subsystem category distribution of the dominant systems identified in the three annotated genomes using the SEED database. Among these subsystems, carbohydrates was the largest subsystem, followed by protein metabolism. The other dominant systems were DNA metabolism and nucleoside and nucleotide. Milk contain high amount of carbohydrates (i.e 5g/100ml) followed by proteins (3.5g/100ml) and fats (1.5/100ml). That is carbohydrate metabolism genes are most common in lactic acid bacteria followed by proteins and fats metabolisms (Buron-Moles *et al*., 2019). The eggnog database, hosted by the European molecular biology laboratory, is a database of orthologous groups and function annotation of the genome. The functional study analysis of the three annotated genomes revealed that the COGs associated with carbohydrate transport and metabolism were dominant followed by those associated with replication, recombination, and repair. The other dominant categories with most COGs associated were amino acid transport and metabolism and translation ribosomal structure and biogenesis, respectively the eggnog database, hosted by the European molecular biology laboratory, is a database of orthologous groups and function annotation of the genome. The functional study analysis of the three annotated genomes revealed that the COGs associated with carbohydrate transport and metabolism were dominant followed by those associated with Replication, recombination, and repair. The other dominant categories with most COGs associated were amino acid transport and metabolism and translation ribosomal structure and biogenesis, respectively.

Funding

This research received no external funding.

Data availability statement

All the data related to this manuscript are included in tables, figures, and supplemental files

Supplementary material

There is supplementary material associated with this article. Access the material online at: https://dx.doi.org/10.17582/journal.pjz/20210708190732

Statement of conflicts of interest

The authors have declared no conflict of interest.

REFERENCES

Adolfsson, O., Meydani, S.N. and Russell, R.M., 2004. Yogurt and gut function. *Am. J. clin. Nutr.*, **80**: 245-256. https://doi.org/10.1093/ajcn/80.2.245

Aldinucci, C., Bellussi, L., Monciatti, G., Passali, G.C., Salerni, L., Passali, D. and Bocci, V., 2002. Effects of dietary yoghurt on immunological and clinical parameters of rhinopathic patients. *Eur. J. clin. Nutr.*, **56**: 1155-1161. https://doi.org/10.1038/sj.ejcn.1601465

Atrih, A., Rekhif, N., Moir, A.J., Lebrihi, A. and Lefebvre, G., 2001. Mode of action, purification and amino acid sequence of plantaricin C19, an anti-Listeria bacteriocin produced by *Lactobacillus plantarum* C19. *Int. J. Fd. Microbiol.*, **68**: 93-104. https://doi.org/10.1016/S0168-1605(01)00482-2

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

Boris, S., Jiménez-Díaz, R., Caso, J.L. and Barbes, C., 2001. Partial characterization of a bacteriocin produced by *Lactobacillus delbrueckii* subsp. *lactis* UO004, an intestinal isolate with probiotic potential. *J. appl. Microbiol.*, **91**: 328-333. https://doi.org/10.1046/j.1365-2672.2001.01403.x

Buron-Moles, G., Chailyan, A., Dolejs, I., Forster, J. and Mikš, M.H., 2019. Uncovering carbohydrate metabolism through a genotype phenotype association study of 56 lactic acid bacteria genomes. *Appl. Microbiol. Biotech.*, **103**: 3135-3152. https://doi.org/10.1007/s00253-019-09701-6

Carver, T., Thomson, N., Bleasby, A., Berriman, M. and Parkhill, J., 2009. DNAPlotter: Circular and linear interactive genome visualization. *Bioinformatics*, **25**: 119-120. https://doi.org/10.1093/bioinformatics/btn578

Conway, P.L., Gorbach, S.L. and Goldin, B.R., 1987. Survival of lactic acid bacteria in the human stomach and adhesion to intestinal cells. *J. Dairy. Sci.*, **70**: 1-12. https://doi.org/10.3168/jds.S0022-0302(87)79974-3

Degnan, F.H., 2008. The US food and drug administration and probiotics: Regulatory categorization. *Clin. Infect. Dis.*, **46**: S133-S136. https://doi.org/10.1086/523324

Fiszman, S.M. and Salvador, A., 1999. Effect of gelatine on the texture of yoghurt and of acid-heat-induced milk gels. *Z. Lebensmittelunters. Forsch. A*, **208**: 100-105. https://doi.org/10.1007/s002170050383

Galperin, M.Y., Makarova, K.S., Wolf, Y.I. and Koonin, E.V., 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucl. Acids Res.*, **43**: D261-269. https://doi.org/10.1093/nar/gku1223

Gezginc, Y., Topcal, F., Comertpay, S. and Akyol, I., 2015. Quantitative analysis of the lactic acid and acetaldehyde produced by *Streptococcus thermophilus* and *Lactobacillus bulgaricus* strains isolated from traditional Turkish yogurts using HPLC. *J. Dairy. Sci.*, **98**: 1426-1434. https://doi.org/10.3168/jds.2014-8447

Ghouri, Y.A., Richards, D.M., Rahimi, E.F., Krill, J.T., Jelinek, K.A. and DuPont, A.W., 2014. Systematic review of randomized controlled trials of probiotics, prebiotics, and synbiotics in inflammatory bowel disease. *Clin. exp. Gastroenterol.*, **7**: 473-487. https://doi.org/10.2147/CEG.S27530

Gurevich, A., Saveliev, V., Vyahhi, N. and Tesler, G., 2013. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics*, **29**: 1072-1075. https://doi.org/10.1093/bioinformatics/btt086

Heller, K.J., 2001. Probiotic bacteria in fermented foods: Product characteristics and starter organisms. *Am. J. clin. Nutr.*, **73**: 374S-379S. https://doi.org/10.1093/ajcn/73.2.374s

Hempel, S., Newberry, S.J., Maher, A.R., Wang, Z., Miles, J.N.V., Shanman, R., Johnsen, B. and Shekelle, P.G., 2012. Probiotics for the prevention and treatment of antibiotic-associated diarrhea: Asystematic review and meta-analysis. *J. Am. med. Assoc.*, **307**: 1959-1969. https://doi.org/10.1001/jama.2012.3507

Hill, C., Guarner, F., Reid, G., Gibson, G.R., Merenstein, D.J., Pot, B., Morelli, L., Canani, R.B., Flint, H.J., Salminen, S., Calder, P.C. and Sanders, M.E., 2014. Expert consensus document. The international scientific association for probiotics and prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat. Rev. Gastroenterol. Hepatol.*, **11**: 506-514. https://doi.org/10.1038/

nrgastro.2014.66

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., Rattei, T., Mende, D.R., Sunagawa, S., Kuhn, M., Jensen, L.J., von Mering, C., Bork, P., 2016. eggNOG 4.5: A hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucl. Acids Res.*, **44**: D286-293. https://doi.org/10.1093/nar/gkv1248

Labayen, I., Forga, L., González, A., Lenoir-Wijnkoop, I., Nutr, R. and Martínez, J.A., 2001. Relationship between lactose digestion, gastrointestinal transit time and symptoms in lactose malabsorbers after dairy consumption. *Aliment. Pharmacol. Ther.*, **15**: 543-549. https://doi.org/10.1046/j.1365-2036.2001.00952.x

Lerebours, E., Ndam, C.N., Lavoine, A., Hellot, M.F., Antoine, J.M. and Colin, R., 1989. Yogurt and fermented-then-pasteurized milk: effects of short-term and long-term ingestion on lactose absorption and mucosal lactase activity in lactase-deficient subjects. *Am. J. clin. Nutr.*, **49**: 823-827. https://doi.org/10.1093/ajcn/49.5.823

Mater, D.D.G., Bretigny, L., Firmesse, O., Flores, M.J., Mogenet, A., Bresson, J.L. and Corthier, G. 2005. *Streptococcus thermophilus* and *Lactobacillus delbrueckii* subsp. *bulgaricus* survive gastrointestinal transit of healthy volunteers consuming yogurt. *FEMS Microbiol. Lett.*, **250**: 185-187. https://doi.org/10.1016/j.femsle.2005.07.006

Mendes-Soares, H., Suzuki, H., Hickey, R.J. and Forney, L.J., 2014. Comparative functional genomics of *Lactobacillus* spp. reveals possible mechanisms for specialization of vaginal *Lactobacilli* to their environment. *J. Bact.*, **196**: 1458-1470. https://doi.org/10.1128/JB.01439-13

Messens, W. and De, V.L., 2002. Inhibitory substances produced by *Lactobacilli* isolated from sourdoughs. A review. *Int. J. Fd. Microbiol.*, **72**: 31-43. https://doi.org/10.1016/S0168-1605(01)00611-0

Meydani, S.N. and Ha, W.K., 2000. Immunologic effects of yogurt. *Am. J. clin. Nutr.*, **71**: 861-872. https://doi.org/10.1093/ajcn/71.4.861

Mohd Adnan, A.F. and Tan, I.K.P., 2007. Isolation of lactic acid bacteria from Malaysian foods and assessment of the isolates for industrial potential. *Bioresour. Technol.*, **98**: 1380-1385. https://doi.org/10.1016/j.biortech.2006.05.034

Ouwehand, A.C., Forssten, S., Hibberd, A.A., Lyra, A. and Stahl, B., 2016. Probiotic approach to prevent antibiotic resistance. *Annls Med.*, **48**: 246-255. https://doi.org/10.3109/07853890.2016.1161232

Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A.R., Xia, F. and Stevens, R., 2014. The Seed and the Rapid Annotation of microbial genomes using subsystems technology (RAST). *Nucl. Acids Res.*, **42**: D206-214. https://doi.org/10.1093/nar/gkt1226

Papadimitriou, K., Zoumpopoulou, G., Foligne, B., Alexandraki, V., Kazou, M., Pot, B. and Tsakalidou, E., 2015. Discovering probiotic microorganisms: *In vitro, in vivo*, genetic and omics approaches. *Front. Microbiol.*, **6**: 58. https://doi.org/10.3389/fmicb.2015.00058

Pintara, A., Jennison, A., Rathnayake, I.U., Mellor, G. and Huygens, F., 2020. Core and accessory genome comparison of Australian and international strains of O157 shiga toxin-producing *Escherichia coli*. *Front. Microbiol.*, **11**: 2162. https://doi.org/10.3389/fmicb.2020.566415

Reid, G., 1999. The scientific basis for probiotic strains of *Lactobacillus*. *Appl. environ. Microbiol.*, **65**: 3763-3766. https://doi.org/10.1128/AEM.65.9.3763-3766.1999

Shao, Y., Gao, S., Guo, H. and Zhang, H., 2014. Influence of culture conditions and preconditioning on survival of *Lactobacillus delbrueckii* subspecies *bulgaricus* ND02 during lyophilization. *J. Dairy Sci.*, **97**: 1270-1280. https://doi.org/10.3168/jds.2013-7536

Simova, E.D., Beshkova, D.M., Angelov, M.P. and Dimitrov, Z.P., 2008. Bacteriocin production by strain *Lactobacillus delbrueckii* ssp. *bulgaricus* BB18 during continuous prefermentation of yogurt starter culture and subsequent batch coagulation of milk. *J. Ind. Microbiol. Biotechnol.*, **35**: 559-567. https://doi.org/10.1007/s10295-008-0317-x

Šmarda, P., Bureš, P., Horová, L., Leitch, I.J., Mucina, L., Pacini, E., Tichý, L., Grulich, V. and Rotreklová, O., 2014. Ecological and evolutionary significance of genomic GC content diversity in monocots. *Proc. Natl. Acad. Sci.*, **111**: E4096-E4102. https://doi.org/10.1073/pnas.1321152111

Smits, T.H.M., 2019. The importance of genome sequence quality to microbial comparative genomics. *BMC Genom.*, **20**: 662. https://doi.org/10.1186/s12864-019-6014-5

Staffolo, M.D., Bertola, N., Martino, M. and Bevilacqua, Y.A., 2004. Influence of dietary fiber addition on sensory and rheological properties of yogurt. *Int. Dairy J.*, **14**: 263-268. https://doi.org/10.1016/j.idairyj.2003.08.004

Sun, Z., Chen, X., Wang, J., Zhao, W., Shao, Y., Guo, Z., Zhang, X., Zhou, Z., Sun, T., Wang, L., Meng, H.,

Zhang, H. and Chen, W., 2011. Complete genome sequence of *Lactobacillus delbrueckii* subsp. *bulgaricus* strain ND02. *J. Bact.*, **193**: 3426-3427. https://doi.org/10.1128/JB.05004-11

Tserovska, L., Stefanova, S. and Yordanova, T., 2002. Identification of lactic acid bacteria isolated from katyk, goat's milk and cheese. *J. Cult. Collect.*, **3**: 48-52.

Van de Water, J., Keen, C.L. and Gershwin, M.E., 1999. The influence of chronic yogurt consumption on immunity. *J. Nutr.*, **129**: 1492S-1495S. https://doi.org/10.1093/jn/129.7.1492S

Vaughan, E.E., de Vries, M.C., Zoetendal, E.G., Ben-Amor, K., Akkermans, A.D.L. and de Vos, W.M., 2002. The intestinal LABs. *Anton. Leeuw. J. Microb.*, **82**: 341-352. https://doi.org/10.1007/978-94-017-2029-8_23