Research Article Open Access

Armenian fermented dairy product "Narine": characterization of the whole-genome sequence of starter bacteria *Lactobacillus helveticus* and potential probiotic marker genes

Karine Chitchyan[1], Arshaluys Verdyan[1], Susanna Hovhannisyan[1], Tigran Soghomonyan[1], Lev Khoyetsyan[1], Anna Mkhitaryan[1], Monika Gogchyan[2], Artur Hambardzumyan[1], Oleg Reva[3], Hrachya Hovhannisyan[1], Valeri Bagiyan[1]

[1]Scientific and Production Center "Armbiotechnology" National Academy of Sciences of the Republic of Armenia, 14 Gyurjyan Str., 0056 Yerevan, Republic of Armenia; [2]Poltava State Medical University, Shevchenko Str. 23, 36000 Poltava, Ukraine; [3]University of Pretoria, Faculty of Natural and Agricultural Sciences, Department of Biochemistry, Genetics and Microbiology, Center for Bioinformatics and Computational Biology, Lynnwood road, Hillcrest, Pretoria, 0002, South Africa

***Corresponding Author:** Artur Hambardzumyan, Scientific and Production Center "Armbiotechnology" National Academy of Sciences of the Republic of Armenia, 14 Gyurjyan Str., 0056 Yerevan, Republic of Armenia.

Submission Date: September 24th, 2025, **Acceptance Date:** November 30th, 2025, **Publication Date:** December 4th, 2025

Please cite this article as: Chitchyan K., Verdyan A., Hovhannisyan S., Soghomonyan T., Khoyetsyan L., Mkhitaryan A., Gogchyan M., Hambardzumyan A., Reva O., Hovhannisyan H., Bagiyan V. Armenian fermented dairy product "Narine": characterization of the whole-genome sequence of starter bacteria Lactobacillus helveticus and potential probiotic marker genes. *Functional Food Science* 2025; 5(12): 703 – 716. DOI: https://doi.org/10.31989/ffs.v5i12.1792

ABSTRACT

Background: The starter monoculture used in "Narine" dairy products was first isolated in 1954 in Armenia by Professor Levon Erzinkyan's team. The strain was later named after the author's granddaughter, Narine, who, according to family legend, was cured of an acute intestinal infection—likely rotavirus-related—through the use of this bacterium. Her condition improved rapidly, and the strain's demonstrated effectiveness prompted further research into its potential for treating intestinal infections. Today, the "Narine" product contains not a commercially formulated mixture but a single strain, with over 60 years of practical use—particularly in fermented milk products and food supplements aimed at both treating and preventing digestive disorders. The strain has shown probiotic properties as well as potential medical applications.

Objective: To elucidate the molecular mechanisms underlying the probiotic and prebiotic properties of this well-known strain, whole-genome sequencing was conducted. Based on the sequence data, potential explanations for its established characteristics were explored.

Methods: Whole-genome sequencing was conducted following Oxford Nanopore Technologies (UK) guidelines, using a Flongle R10.4.1 flow cell and the ligation sequencing kit SQK-LSK114 as specified in the *g-dna-by-ligation-sqk-lsk114* protocol. High-molecular-weight, high-purity DNA was extracted via the traditional phenol-chloroform method. Raw sequencing data were processed with Flye (version 2.9.1) and annotated using Prokka (version 1.14.5), a bacterial genome annotation tool.

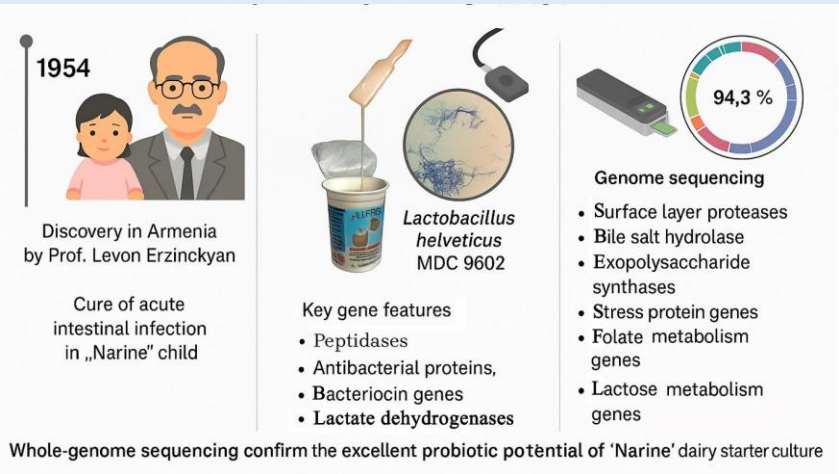
Results: Based on NGS sequencing, the "Narine" starter culture was identified as *Lactobacillus helveticus*. At this stage, the genome of our strain comprises 1,923,890 nucleotides across 490 contigs, without circular structures, representing 94.3% of the reference strain genome. Preliminary genome analysis reveals several genes associated with known probiotic characteristics of the strain. These include surface layer proteases (2 genes), cell division proteins (6 genes), bile salt hydrolase, biofilm regulatory protein A, exopolysaccharide synthases (3 genes), lipoteichoic acid synthases (3 genes), GroEL and Clp stress protein genes (10 genes), folate metabolism genes (11 genes), lactose metabolism genes (18 genes), peptidases (36 genes), and others.

Novelty: The complete genome of the starter culture of the 'Narine' dairy product, which has been used as a prophylactic nutritional supplement for newborns in the Soviet Union since the 1960s, has been characterized for the first time. The probiotic properties of the starter culture have been demonstrated.

Conclusion: Preliminary studies of the newly sequenced *Lactobacillus helveticus* genome demonstrate the excellent probiotic potential of the well-known "Narine" dairy product's starter culture.

Keywords: Dairy product, "Narine", *Lactobacillus helveticus*, whole bacterial genome, Oxford Nanopore Technologies, probiotic

Graphical Abstract: Armenian fermented dairy product "Narine": characterization of the whole-genome sequence of starter bacteria *Lactobacillus helveticus* and potential probiotic marker genes



INTRODUCTION

A joint FAO/WHO expert consultation on the health and nutritional properties of live lactic acid bacteria powders was held from 1 to 4 October 2001 at the American Cordoba Park Hotel in Córdoba, Argentina. The consultation, which was the first meeting of this group, focused on assessing the available scientific evidence on the properties, functionality, benefits, safety, and nutritional characteristics of probiotic foods. Probiotics are gram-positive bacteria, primarily belonging to two genera: *Lactobacillus* and *Bifidobacterium* [1]. Now the World Health Organization defines probiotics as "live microorganisms which when administered in adequate amounts confer a health benefit on the host" (FAO/WHO). The World Health Organization now defines probiotics as "live microorganisms which, when administered in adequate amounts, confer a health benefit on the host".

Health professionals are increasingly promoting the beneficial effects of foods with added live microorganisms (probiotics) on human health, particularly dairy products, for children and other high-risk groups. It has been reported that these probiotics may play important roles in immunological, digestive, and respiratory functions, and may have a significant impact on alleviating infectious diseases in children [1]. The discovery of probiotic properties in microorganisms and the interpretation of their mechanisms are complicated by their multifactorial nature. In this regard, new opportunities are emerging as a result of whole-genome sequencing of probiotic microorganisms. Identifying known genes within the genome of a probiotic organism enables us to better understand its various probiotic characteristics, despite the presence of a significant percentage of hypothetical genes.

The complete genome sequence, combined with the demonstration of a low frequency of undesirable characteristics – such as plasmids, prophages, antibiotic resistance genes, and virulence factors – and frequent bacteriocin production, facilitates the identification of the strain and its potential probiotic properties. These traits support the classification of this species as a good candidate for use as a probiotic [2,3]. Novadeys Next Generation Sequencing (NGS) is a common procedure for whole bacterial genome sequencing [4]. However, sequencing the complete genome of probiotic bacteria is more successful when long-read methods are used, such as PacBio [5] or Oxford Nanopore [6].

The renowned Armenian functional food for newborns, "Narine," was developed based on Professor Yerzinkyan's starter cultures, Er-1 and Er-2, which he isolated from newborn meconium in the late 1940s. Later, the strain Er-2 was identified as *Lactobacillus acidophilus* n.v. Ep317/402 using classical classification methods and received the USSR Author's Certificate No. N163573 in 1964 (https://narine.okis.ru/).

In his 1965 book, the scientist wrote: *"The local strains of acidophilic lactic acid bacteria isolated by us, in terms of cultural, morpho-physiological, biochemical, and organoleptic properties, differ favorably from those previously described in the scientific literature in their identifiers of acidophilic bacterial microorganisms. The main features of our strains from the Er-1 and Er-2 groups are their high antimicrobial properties, resistance to phenol (0.5%), chemotherapeutic agents (phthalazole resistance up to 1%), and antibiotics (synthomycin resistance up to 0.003%), which is of particular importance for medicine and veterinary science"* [7].

The product is named after Professor Yerzinkyan's granddaughter, "**Narine"**, who became the first child to receive this fermented milk as a substitute for breast milk. Yerzinkyan also discusses the starter culture in another book published in 1971 [8]. The majority of Yerzinkyan's works, including more than 200 publications and 20 patents – many of which concern **"Narine"** acidophilic milk or its starter cultures – are summarized in the previously mentioned citation (https://narine.okis.ru/).

Antimicrobial peptides (AMPs) have garnered attention as alternatives to chemical food preservatives and widely used antibiotics. The *Lactobacillus acidophilus* n.v. Er 317/402 strain used in Narine produces a small AMP with a molecular weight of 1.1 kDa. This AMP is highly heat-stable (remaining active after 90 minutes at 130°C), functions over a wide pH range, and is sensitive to proteolytic enzymes such as trypsin, pepsin, and proteinase K. The peptide exhibits broad-spectrum activity against both Gram-positive and Gram-negative pathogens, including some that the World Health Organization classifies as hazardous infections. Its preliminary amino acid sequence is estimated as Asn–Val–Gly–Val–Leu–X_1–Pro–Pro–X_2–Leu–Val, where X_1 can be Leu or Asn, and X_2 can be Met or Pro [9].

The functional food **"Narine"** is a fermented milk product or a freeze-dried live culture of the strain *Lactobacillus acidophilus* MDC 9602. Recent studies have shown that **"Narine"** lactic acid bacteria are resistant to a range of antibiotics and chemotherapeutic agents; possess high antagonistic activity against a broad spectrum of pathogenic and conditionally pathogenic microorganisms

Recent studies have shown that **"Narine"** starter lactic acid bacteria are resistant to a number of antibiotics and chemotherapeutic agents, and exhibit high antagonistic activity against a broad spectrum of pathogenic and conditionally pathogenic microorganisms (including agents of dysentery, enteric fever, *salmonellosis*, pathogenic *Escherichia coli*, *streptococci*, *staphylococci*, and *Proteus*); help restore healthy normal intestinal microflora and improve its motor functions; and promote the synthesis of B, C, E, and P group vitamins, as well as folic acid and biotin in the body [10,11].

Based on polymerase chain reaction (PCR) amplification and sequencing of 16S rRNA gene fragments and the RNA polymerase rpoA gene [NCBI GenBank accession numbers HQ379170 and HQ379179, Pashayan MM, Hovhannisyan HG, 2010], along with a series of morphological, physiological, and biochemical tests, the starter culture of **"Narine"** was renamed *Lactobacillus helveticus* MDC 9602. In vitro evaluation of the adaptive and probiotic properties of *L. helveticus* MDC 9602, the Armenian functional sour milk starter used in "Narine," demonstrated its broad potential in the treatment of gastrointestinal disorders and in lowering blood pressure. The adaptive properties of MDC 9602, which are essential for bacterial survival in the gastrointestinal tract, include high acid tolerance, strong autoaggregation, and hydrophobicity. *L. helveticus* MDC 9602 produces significant amounts of lactic acid and exhibits antagonistic effects against pathogenic and related bacteria [12]. The effect of lemon molasses alcoholic extract on the growth of *Lactobacillus helveticus* MDC 9602 and the organoleptic and antioxidant properties of the fermented milk [13], as well as the potential to improve the functional and sensory qualities of the fermented dairy drink **"Narine"** using raw apricot gum [14].

The genomes of five L. helveticus strains were sequenced to completion and compared with those of other genomically characterized lactobacilli. *L. helveticus* has the potential to produce peptides with a biological function, such as angiotensin converting enzyme inhibitory activity, in fermented dairy products, demonstrating the therapeutic value of this species. A most intriguing feature of the genome of *L. helveticus* is the remarkable similarity in gene content with many intestinal lactobacilli. Comparative genomics has enabled the identification of key gene sets that facilitate diverse lifestyles, including adaptation to food matrices and the gastrointestinal tract. As genome sequences and functional genomic information continue to explode, key features of the genomes of L. helveticus strains are being discovered, answering many questions but also raising many new ones [15]. Up-to-date whole-genome

sequencing of *Lactobacillus helveticus* is an urgent task [16,17].

Long-read Oxford Nanopore sequencing is now widely used for microbial genome sequencing and enables the recovery of highly contiguous microbial genomes from isolates or metagenomes. It has been shown that Oxford Nanopore R10.4 can generate near-finished microbial genomes from isolates or metagenomes without the need for short-read or reference-based polishing [18]. ONT sequencing is increasingly used in microbial genomics due to its ability to perform long-read, real-time sequencing. Its performance and applicability for genome reconstruction across various bacterial species have been demonstrated [19].

This study aimed to perform whole-genome sequencing of the well-known functional food starter culture 'Narine' using Oxford Nanopore Technology (ONT) next-generation sequencing (NGS) to confirm the strain's accurate classification, and to use genomic and proteomic data to identify its key probiotic and prebiotic characteristics.

METHODS

Strain cultivation: *Lactobacillus helveticus* MDC 9602 strain biomass for DNA extraction was obtained by cultivation in a laboratory-developed growth medium optimized by our research group. The medium composition per 1000 mL was as follows: skim milk powder–5 g, peptone–10 g, yeast extract–5 g, glucose–20 g, Tween 80–1 g, KH_2PO_4–3 g, sodium acetate–4 g, sodium citrate–2 g, $MgSO_4$–0.2 g, $MnSO_4$–0.05 g; pH was adjusted to 6.5 before sterilization. Cultivation was carried out at 37°C for 18 h under microaerophilic conditions prior to DNA isolation.

DNA isolation: High-molecular-weight, high-purity DNA was isolated using a classic phenol-chloroform extraction method [20].

Sequencing: Whole-genome sequencing was performed following the recommendations of Oxford Nanopore Technologies (UK), using a Flongle R10.4.1 flow cell and the ligation sequencing kit SQK-LSK114, according to the protocol specified in the g-dna-by-ligation-sqk-lsk114-document (document: Flongle-en-GDE_9161_v114_revY_30Jan2025.pdf).

Sequencing data processing: Bacterial genome assembly was performed using Flye (version 2.9.1), a de novo genome assembler designed for long-read sequencing data [21,22]. It is optimized for assembling high-error-rate reads into high-quality, polished contiguous sequences (contigs). The software is available at: https://github.com/mikolmogorov/Flye /releases/tag/2.9.1.

Bacterial genome annotation: Bacterial genome annotation was performed using Prokka (version 1.14.5), a prokaryotic genome annotator that provides high-quality, standardized annotations [23]. The software is available at: https://github.com/tseemann/prokka/releases/tag/v1.14.5.

Bacteriocin gene prediction: Bacteriocin-producing genes were predicted using BAGEL4 (http://bagel4.molgenrug.nl/) (accessed on 20 October 2025).

RESULTS AND DISCUSSION

Assembly of "Narine" starter culture genome: The Flye assembly of the "Narine" starter culture genome resulted in a total length of 1,923,890 base pairs (bp) across 490 contigs, with an average coverage of 96x (see flye.log in Mendeley Data). The longest contig is 19,332 nucleotides, and the N50 statistic – representing the contig length at which 50% of the genome is contained in contigs of this size or larger – is 4,401 bp. The genome contains 2,218 coding sequences (CDSs), 4 rRNAs, 3 repeat regions, 40 tRNAs and 1 tmRNA (see PROKKA_07252025.txt in Mendeley Data). The

chromosome schematic is presented in Fig. S1 of supplementary materials in Mendeley Data. No circular structures were detected; however, three plasmids were identified: in contig 483 (9,857 nucleotides, 2 copies, Plasmid 1), contig 1 (3,594 nucleotides, 4 copies, Plasmid 2), and contig 118 (1,555 nucleotides, 7 copies, Plasmid 3). Refer to Supplementary Table S1 and Supplementary Figures S2–S4 in Mendeley Data.

Classification of "Narine" starter culture strain: The BLAST analysis of 1,560 nucleotides of the 16S rRNA gene from the "Narine" starter culture (see Supplementary Fig. S5 in Mendeley Data), located in contig 458 (17,361 nucleotides), revealed 100% identity with the corresponding gene of *Lactobacillus helveticus* strain BIM B-461 G (accession CP097886).

The BLAST analysis of 4,314 nucleotides from contig 235, which contains the DnaA gene (see Supplementary Fig. S2 in Mendeley Data), revealed 99.98% identity with the *Lactobacillus helveticus* isolate MGYG-HGUT-02384 genome assembly (accession LR698986). Based on this, we classified the "Narine" starter culture strain as *Lactobacillus helveticus*, designated as MDC 9602. When compared to the genome of *L. helveticus* isolate MGYG-HGUT-02384, the sequenced length of *L. helveticus* MDC 9602 covers approximately 94.3% of the reference genome. Future research will focus on achieving 100% genome resolution for our strain.

Characterisation of L. helveticus MDC 9602 proteome: The following proteins of L. helveticus, along with their contig deposition, CDS locations, and coding gene chain positions, are presented in Table 1: surface layer proteins, biofilm regulatory protein, bile salt hydrolase, cell division proteins, and lipoteichoic acid synthases.

Table 1. Surface-layer, biofilm regulatory, bile salt hydrolase, cell division, and lipoteichoic acid synthesis proteins of *Lactobacillus helveticus.*

Protein	Contig[1]	CDS[2]	Chain[3]
S-layer protein SlpH_1	362	56-430	+
S-layer protein SlpH_2	426	1133-2320	+
Biofilm regulatory protein A	434	5961-7070	+
Bile salt hydrolase BshA	242	476-1591	-
Putative cell division protein WhiA	326	997-1932	-
Cell division protein FtsL	436	1789-2151	+
Cell division protein SepF	10	2029-2466	-
Cell division protein FtsZ	10	2484-3803	-
Cell division protein FtsA	10	3818-4717	-
Cell division protein DivIB	236	659-1516	-
Lipoteichoic acid synthase 1	14	306-1709	+
Lipoteichoic acid synthase 1	97	617-1210	+
Lipoteichoic acid synthase 2	125	80-2233	-

(1) The contig number on which the protein is located; (2) the range of nucleotides on the specified contig containing the protein-coding sequence (CDS); (3) the chain of contigs where the protein is located (positive strand or negative complement).

The gene groups identified in *L. helveticus* MDC 9602 underscore key probiotic traits. S-layer proteins in probiotics often mediate adhesion to intestinal epithelial cells and facilitate binding to extracellular matrix components such as collagen, laminin, and fibronectin. They also play a vital role in biofilm formation and colonization of host mucosal surfaces. Overall, S-layer proteins contribute significantly to the probiotic's ability

to survive in the gastrointestinal environment, establish stable colonization, and interact beneficially with the host [24]. Biofilm regulatory proteins are essential for effective colonization, resilience, and functional activity of probiotics in the gut. They support the formation of stable microbial communities, enhance resistance to environmental stresses, and promote health benefits to the host. Bile salt hydrolases are crucial enzymes that enable probiotic bacteria to withstand bile toxicity, thereby improving their survival and functionality within the gastrointestinal tract. These enzymes also contribute to cholesterol reduction and modulation of the gut microbiota, supporting overall host health [25].

Cell division proteins, represented by six homologs in the *Lactobacillus helveticus* genome, are crucial for maintaining stable bacterial growth and successful colonization within the gastrointestinal environment. They support colonization by ensuring efficient cell division, maintaining the stability of the probiotic population, adapting to environmental stresses, and facilitating effective interaction with the host. These proteins are fundamental to the proliferation, persistence, and functional stability of probiotics, thereby directly influencing their efficacy and ability to confer health benefits.

The probiotic significance of lipoteichoic acid synthases (LTA synthases), exemplified by three homologous genes within the *Lactobacillus helveticus* genome, encompasses a range of vital physiological functions. These include, but are not limited to, immunomodulation, adhesion to intestinal epithelium, colonization stability, maintenance of cell wall integrity, and overall probiotic efficacy. LTA synthases are crucial enzymes responsible for the biosynthesis of lipoteichoic acids, which play essential roles in bacterial cell wall architecture and function. By mediating LTA production, these enzymes influence bacterial adhesion, resilience within the gastrointestinal environment, and the capacity to modulate host immune responses, thereby underpinning the probiotic potential of *L. helveticus* and other Gram-positive probiotic bacteria [26].

The *Lactobacillus helveticus* genome encodes chaperonins, ATP-dependent Clp proteases, exopolysaccharide synthases, and lactose metabolism-associated proteins – key determinants of probiotic functionality – which are summarized in Table 2.

Table 2. Chaperonins, ATP-dependent Clp proteases, exopolysacharide synthases, and lactose metabolism enzymes of *Lactobacillus helveticus.*

Protein	Contig[1]	CDS[1]	Chain[1]
10 kDa chaperonin GroS_1	107	1623-1907	+
10 kDa chaperonin GroS_2	357	1300-1584	-
33 kDa chaperonin HslO	470	1048-1938	-
ATP-dependent Clp protease proteolytic subunit ClpP	108	21-605	+
ATP-dependent Clp protease ATP-binding subunit ClpX	12	47-604	+
ATP-dependent protease ATPase subunit ClpY	174	6593-7996	+
ATP-dependent Clp protease ATP-binding subunit ClpE_1	280	95-2149	+
ATP-dependent Clp protease ATP-binding subunit ClpE_2	31	4820-5533	+
ATP-dependent Clp protease ATP-binding subunit ClpE_3	98	3431-5239	+
ATP-dependent Clp protease ATP-binding subunit ClpC	422	2351-4831	-
Putative sugar transferase EpsL_1	165	1158-1808	+
Putative sugar transferase EpsL_2	242	1723-2124	-
Putative glycosyltransferase EpsH	379	463-1278	+
6-Pphospho-beta-galactosidase LacG	111	1229-1810	-

Protein	Contig[1]	CDS[1]	Chain[1]
HTH-type transcriptional regulator LacR	148	846-1853	-
Protein LacX_1, plasmid	174	8158-9057	+
Protein LacX_2, plasmid	7	74-589	+
Lactose permease LacS_1	300	347-676	+
Lactose permease LacS_2	300	695-1819	+
Lactose permease LacS_3	300	2083-3885	+
Beta-galactosidase large subunit LacL_1	332	39-659	+
Beta-galactosidase large subunit LacL_2	380	45-1931	+
Beta-galactosidase small subunit LacM_1	332	643-1599	+
Beta-galactosidase small subunit LacM_2	380	1915-2223	+
Lactose transport system permease protein LacF	453	12475-13341	+
Galactoside O-acetyltransferase LacA	461	3166-3780	+
Galactokinase galK_1 GalK_1	407	3236-4144	+
Galactokinase galK_1 GalK_2	418	3958-4782	-
Galactose-1-phosphate uridylyltransferase GalT	418	2473-3936	-
UDP-glucose 4-epimerase GalE_1	206	37-243	+
UDP-glucose 4-epimerase GalE_2	332	1704-2369	+

(1) See explanations under Table 1.

We observed two copies of *GroS* and one *Hsl* chaperonin gene in the *Lactobacillus helveticus* genome. Chaperonins, such as GroEL and GroES, are specialized proteins that assist in the proper folding of newly synthesized or stress-denatured proteins. In probiotics, chaperonins enhance survival under harsh gastrointestinal conditions by preventing protein misfolding caused by bile acids, acidity, and other stressors. They also support the maintenance of cellular functions essential for colonization and interaction with the host [27].

The seven ATP-dependent Clp proteases of *Lactobacillus helveticus* are also listed in Table 2. These proteases (such as ClpXP, ClpAP, etc.) are responsible for degrading misfolded, damaged, or regulatory proteins, thereby maintaining protein quality control. In probiotics, Clp proteases contribute to stress resistance, oxidative protection, and adaptation to the intestinal environment. They also play a role in regulating gene expression related to stress responses, colonization, and persistence within the host [28].

As shown in Table 2, the genome of *Lactobacillus helviticus* includes three exopolysaccharide synthase (EPS) genes. EPS synthesis genes are essential for the probiotic qualities of bacteria, and their significance is associated with enhanced adhesion and colonization, biofilm formation, immunomodulation, protection against environmental stresses, and various health benefits [29].

Finally, 18 gene products related to lactose metabolism in the *Lactobacillus helveticus* genome are listed in Table 2. Proteins involved in lactose metabolism are essential for the probiotic characteristics of bacteria, contributing to their survival and growth in the gut, enhancing colonization, alleviating lactose intolerance symptoms, and providing metabolic and health benefits [30].

The genes of *Lactobacillus helveticus* associated with folate metabolism, including six folate synthesis genes, are listed in Table 3.

Table 3. Enzymes of *Lactobacillus helveticus* associated with folate metabolism

Protein	Contig[1]	CDS[1]	Chain[1]
Folylpolyglutamate synthase	19	631-1968	+
Dihydrofolate reductase (folA)	163	2239-2652	+
Folate transporter FolT	130	572-1090	+
GTP cyclohydrolase 1 type 2	59	1255-2052	+
GTP cyclohydrolase 1, folE	19	144-647	+
Dihydrofolate synth./folylpolyglutamate synth.	156	611-1234	+
MTHF-tRNA-methyltransferase TrmFO	174	3825-5141	+
5,10-MTHF reductase	30	3160-3825	-
Formate-tetrahydrofolate ligase 1	305	920-2419	-
Formate-tetrahydrofolate ligase 1	41	3945-5606	-
Formate-tetrahydrofolate ligase	39	5402-5635	+

(1) See explanations under Table 1.

Folate metabolism is highly relevant to the probiotic characteristics of bacteria because folate (vitamin B9) plays a crucial role in both microbial and host health. Its significance includes supporting bacterial growth and survival, enabling the production of bioactive compounds, modulating host health, competing with pathogens, and maintaining gut microbiota homeostasis [31]. Given that folate is critically important for the health of pregnant women and newborns – due to its essential roles in fetal development and maternal well-being, including neural tube formation, cell growth and division, prevention of birth defects, maternal health, and postnatal benefits – we can conclude that the probiotic strain-based product "Narine" offers valuable health-promoting properties related to folate synthesis.

Four bacteriocins mined from the *Lactobacillus helveticus* genome using BAGEL4 and annotated as hypothetical by Prokka are presented in Table 4.

Table 4. BAGEL4 predicted bacteriocin genes in the *Lactobacillus helveticus* genome.

Protein predicted by BAGEL4	Protein BLAST[1]	% of homology[2]	Contig[3]	CDS[3]	Chain[3]
Helveticin-J	Helveticin J family class III bacteriocin, partial [*Lactobacillus amylovorus*]	92.56	110	1312-2274	-
Helveticin-J	Helveticin J family class III bacteriocin [Lactobacillus helveticus]	97.23	27	7413-8162	+
Enterolysin_A	M23 family metallopeptidase [Lactobacillus helveticus]	100.00	303	2866-3411	+
Enterolysin_A	M23 family metallopeptidase [Lactobacillus helveticus]	97.71	318	4188-4844	+

(1) Nierest homologous protein after Protein BLAST in NCBI database, (2) homology percentage after BLAST, (3) See explanations under Table 1.

Helveticin J is a type of bacteriocin produced by certain strains of *Lactobacillus helveticus*. As a class III bacteriocin, it is characterized by being a large, heat-labile protein with antimicrobial activity, and its mechanism of action has not yet been fully established. Regarding Enterolysin A, the M23 family metallopeptidases are a group of zinc-dependent enzymes that hydrolyze peptide bonds in bacterial cell wall peptidoglycan, resulting in cell lysis.

As presented in Supplementary Table S2 (Mendeley Data), the genome of *Lactobacillus helveticus* contains at least 36 different peptidase genes. In lactic acid bacteria (LAB), certain peptidases can influence probiotic properties by directly contributing to: (i) protein degradation through the release of bioactive peptides

with health benefits (e.g., antimicrobial, immunomodulatory, antihypertensive); (ii) stress tolerance and survival, as peptide and amino acid recycling supports growth under gut conditions; and (iii) adhesion and interaction with the host via modification of surface proteins or production of signaling peptides [32]. For instance, aminopeptidase YpdF is a member of the M24 family of metallopeptidases, which are found across bacteria, archaea, and eukaryotes, with functions ranging from bacterial cell wall turnover to protein maturation in various organisms. Therefore, a high abundance of peptidases may serve as evidence of the strong probiotic potential of the studied strain.

Lactic acid synthesis by LAB is essential for their probiotic functionality because it directly contributes to pathogen suppression, maintenance of gut health, and immune modulation. These properties underpin many of the health benefits associated with the use of lactic acid bacteria as probiotics. Enzymes of *Lactobacillus helveticus* associated with lactic acid synthesis are listed in Supplementary Table S3 (Mendeley Data).

Genes of L. helveticus MDC 9602 harbored on plasmid-like structures: The protein genes of *L. helveticus*, which are harbored on plasmid-like structures (see Table 5), were mainly annotated by Prokka as hypothetical. To investigate their potential biological functions, we performed additional Protein BLAST analyses of their respective amino acid sequences against the latest NCBI database.

The largest contig, number 483 (9,857 bp), contains 8 coding sequences (CDSs): seven of which, as annotated by Prokka (version 1.14.5), are hypothetical proteins. The eighth CDS was identified as a putative iron export ATP-binding protein FetA. The fourth CDS remained hypothetical after BLAST analyses, while the other six were annotated as follows:

Table 5. Genes of *L. helveticus* MDC 9602 harbored on plasmid-like structures.

Protein	Protein BLAST[1]	% of homology[1]	Contig[1]	CDS[1]	Chain[1]
Hypothetical protein	Mbeg1-like protein	95.24	483	251-379	-
Hypothetical protein	MgtC/SapB family protein	97.86	483	463-1167	-
Hypothetical protein	Eco57I restriction-modification methylase domain-containing protein	89.40	483	1704-6110	+
Hypothetical protein	MULTISPECIES: hypothetical protein [*Lactobacillus*]	99.10	483	6110-7108	+
Hypothetical protein	DNA methyltransferase [*Lactobacillus kefiranofaciens*]	85.40	483	7148-7828	+
Hypothetical protein	DUF3278 domain-containing protein [*Lactobacillus helveticus*]	97.87	483	8052-8195	-
Hypothetical protein	helix-turn-helix transcriptional regulator [*Lactobacillus helveticus*]	95.45	483	8192-8392	-
Putative iron export ATP-binding protein FetA			483	8560-9198	+
Hypothetical protein	protein YbcW [*Escherichia coli*]	98.53	1	601-807	-
Putative protein YdfO	head protein [*Escherichia phage Lambda*]	100.00	1	1051-1503	+
Hypothetical protein	serum resistance lipoprotein Bor [*Escherichia coli*]	98.97	1	1793-2086	+
Hypothetical protein	lysis protein [*Escherichia coli]*	99.15	1	2209-2577	-
Hypothetical protein	glycoside hydrolase family 24 protein [*Escherichia coli*]	98.10	1	2574-3050	-
Hypothetical protein	TPA_asm: phage holin, lambda family [*Escherichia coli*]	100.00	1	3190-3357	-
IS200/IS605 family transposase ISLhe65			118	228-1418	-

(1) See explanations under Table 4.

- **Mbeg1-like protein:** These proteins are generally associated with a family involved in bacterial cell wall synthesis, motility, or gene regulation. In probiotic bacteria, Mbeg1-like proteins may facilitate adhesion to host tissues, biofilm formation, and resistance to environmental stresses, contributing to probiotic efficacy.
- **MgtC/SapB family protein:** These proteins are crucial for bacterial survival under stress conditions. They influence virulence in pathogenic bacteria and may enhance resilience and colonization in probiotic strains. Their functions can vary among species but generally support adaptation within the host gut.
- **Eco57I restriction-modification methylase domain-containing protein:** This enzyme is part of a bacterial restriction-modification system, methylating specific DNA sequences to protect the host genome from phage infection and foreign DNA.
- **DNA methyltransferase (*Lactobacillus kefiranofaciens*):** This enzyme functions primarily to methylate bacterial DNA as part of a restriction-modification system, maintaining genome stability, protecting against foreign genetic elements, and potentially influencing probiotic performance.
- **DUF3278 domain-containing protein (*Lactobacillus helveticus*):** Although its exact function remains uncharacterized, its conservation suggests a role in probiotic survival, adaptation, or host interaction, warranting further functional investigation.
- **Helix-turn-helix transcriptional regulator (*Lactobacillus helveticus*):** These regulators are fundamental components of gene regulation networks that enable the bacteria to respond adaptively to environmental stresses, thereby influencing probiotic properties and interactions with the host.

The involvement of all these proteins highlights their significant contribution to the probiotic characteristics and adaptive capacity of *Lactobacillus helveticus* MDC 9602.

The contig number 1 (3,594 bp) contains 6 CDSs: five of which, as annotated by Prokka (version 1.14.5), are hypothetical proteins. The second CDS was annotated as a head protein from the *Escherichia phage Lambda*. The five hypothetical proteins, based on BLAST analyses, were annotated as follows:

- **YbcW (Escherichia coli):** Currently classified as a hypothetical or putative protein with undefined functions. Further experimental validation is needed to clarify its role in bacterial physiology, stress response, or other cellular processes.
- **Serum resistance lipoprotein Bor:** In *E. coli*, Bor functions as an immune evasion factor, facilitating bacterial survival in the bloodstream by conferring resistance to complement-mediated killing. Its expression is associated with increased virulence and pathogenic potential.
- **Lysis proteins:** In *E. coli*, these proteins are primarily involved in degrading the bacterial cell wall during phage-induced lysis, playing a crucial role in the phage life cycle and offering potential applications in antibacterial strategies.
- **Glycoside hydrolase family 24 proteins:** These are peptidoglycan hydrolases that catalyze cleavage within cell wall components, playing essential roles in bacterial cell growth, division, and maintenance of cell envelope integrity.
- **TPA_asm: phage holin, lambda family:** This membrane-associated protein forms pores to facilitate cell lysis during phage reproduction, playing a pivotal role in the phage lifecycle and bacterial cell death.

In summary, this plasmid-like contig appears to encode proteins associated with both *E. coli* and the *E. coli* phage lambda family. These proteins may participate in targeting competing bacteria within the gut microbiota, potentially demonstrating probiotic characteristics through bacterial antagonism and antimicrobial activity [33–35].

The contig number 118 (1555 bp) contains only one CDS, which encodes IS200/IS605 family transposase ISLhe65. IS200/IS605 family transposase ISLhe65 is a mobile genetic element that facilitates genome rearrangements in bacteria. Its activity can affect the

genetic diversity, adaptability, and functional traits of Lactobacillus strains and other bacteria, including those with probiotic potential.

CONCLUSION

The Flye assembly of the "Narine" starter culture genome yielded a total length of 1,923,890 bp distributed across 490 contigs. This genome size is typical for *Lactobacillus helveticus* strains.

Based on BLAST analysis of contig 235, which harbored the *dnaA* gene, we classified the "Narine" starter culture strain as *Lactobacillus helveticus*. The assembly covers approximately 94.3% of the known reference genomes for *L. helveticus*. Future research will focus on achieving a complete, 100% genome resolution of the *L. helveticus* MDC 9602 strain.

The genome analysis reveals several genes associated with the strain's known probiotic characteristics. These include: surface layer proteases (2 genes), cell division proteins (6 genes), bile salt hydrolase, biofilm regulatory protein A, exopolysaccharide synthases (3 genes), lipoteichoic acid synthases (3 genes), GroEL and Clp stress protein genes (10 genes), folate metabolism genes (11 genes), lactose metabolism genes (18 genes), bacteriocins (4 genes), peptidases (36 genes), lactate dehydrogenases (4 genes) and others.

The functions of genes located within the plasmid-like structures of *Lactobacillus helveticus* MDC 9602 suggest they significantly contribute to probiotic traits and adaptive capacity. These genes may target competing bacteria within the gut microbiota and indicate the presence of mobile genetic elements that can influence genetic diversity, adaptability, and functional traits of *Lactobacillus* strains and other bacteria. This genomic mobility can potentially enhance probiotic efficacy and facilitate the acquisition of beneficial features.

In summary, based on the above, we can conclude that, for the first time, the complete genome of the starter culture of the 'Narine' dairy product has been characterized. The excellent probiotic properties of the starter culture have been demonstrated.

Abbreviations: AMPs: antimicrobial peptides; CDSs: coding sequences; Clp: Caseinolytic protease; FAO: Food and Agriculture Organization; LAB: lactic acid bacteria; LSK: Ligation Sequencing Kit; MDC: Microbial Depository Center; NGS: Next Generation Sequencing; ONT: Oxford Nanopore Technology; PCR: polymerase chain reaction; SQK: Sequencing Kit; WHO: World Health Organization

Data Availability Statement: The supplementary materials (SupplementaryMaterial_Narine.docx) and genome data files (flye.log, PROKKA_07252025.faa, PROKKA_07252025.ffn, PROKKA_07252025.gbk, and PROKKA_07252025.txt) for this article are available at https://data.mendeley.com/datasets/r5s89w8g7k/3 under the "Whole Genome of *Lactobacillus helveticus* MDC 9602" project of Mendeley Data.

Competing Interests: The authors have no financial interests or conflicts of interest.

Author's Contributions: Conceptualization: AH, OR, TS, LKh; Methodology: AH, KCh, AV, TS, SH, MG, and AM; Software: AH, OR, LKh., TS, AM, MG; Validation: LKh, TS, AM; Resources: AH, HH, VB; Writing review & editing: AH, OR, HH, VB. All authors have read and agreed to the published version of the manuscript.

Acknowledgements and Funding: This research was supported by the Committee of Higher Education and Science of the Republic of Armenia, within the framework of research project № 21AG-2I090.

REFERENCES

1. Anonyme. Report of a joint FAO/WHO Expert Consultation. Health and nutritional properties of probiotics in food includion powder milk with live lactic acid bacteria. Cordoba, Argentina, October 2001. n.d. https://openknowledge.fao.org/items/db384295-64d9-47e2-b65b-3c918efc5140. Retrieved on December 4th, 2025

2. Torres-Miranda A, Melis-Arcos F, Garrido D. Characterization and Identification of Probiotic Features in Lacticaseibacillus Paracasei Using a Comparative Genomic Analysis Approach. *Probiotics Antimicrob Proteins* 2022;14(6):1211–24. DOI: https://doi.org/10.1007/s12602-022-09999-1
3. Barreto Pinilla CM, Guzman Escudero F, Torres Silva E Alves A, Spadoti LM, Brandelli A. Draft Genome Sequence and Comparative Genome Analysis Reveal Potential Functional Properties in Lacticaseibacillus paracasei ItalPN16. *Curr Microbiol* 2023;80(12):399. DOI: https://doi.org/10.1007/s00284-023-03515-6
4. Shokralla S, Spall JL, Gibson JF, Hajibabaei M. Next-generation sequencing technologies for environmental DNA research. *Mol Ecol* 2012;21(8):1794–805. DOI: https://doi.org/10.1111/j.1365-294X.2012.05538.x
5. Rhoads A, Au KF. PacBio Sequencing and its Applications. *Genomics Proteomics Bioinformatics* 2015;13(5):278–89. DOI: https://doi.org/10.1016/j.gpb.2015.08.002
6. Wick RR, Judd LM, Gorrie CL, Holt KE. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genomics* 2017;3(10). DOI: https://doi.org/10.1099/mgen.0.000132
7. Yerzinkyan Levon A. *Preparation and use of medicinal acidophilic milk and milk "Narine". [Published in Russian]*. Yerevan: Publishing House of the Academy of Sciences of the Armenian SSR. 1965
8. Yerzinkyan Levon A. *Biological peculiarities of some races of lactic acid bacteria*. Yerevan: Publishing House of the Academy of Sciences of the Armenian SSR. 1971
9. Mkrtchyan H, Gibbons S, Heidelberger S, Zloh M, Limaki HK. Purification, characterisation and identification of acidocin LCHV, an antimicrobial peptide produced by Lactobacillus acidophilus n.v. Er 317/402 strain Narine. *Int J Antimicrob Agents* 2010;35(3):255–60. DOI: https://doi.org/10.1016/j.ijantimicag.2009.11.017
10. Pepoyan AZ, Balayan MH, Manvelyan AM, Mamikonyan V, Isajanyan M, Tsaturyan VV et al. *Lactobacillus acidophilus* INMIA 9602 Er-2 strain 317/402 probiotic regulates growth of commensal *Escherichia coli* in gut microbiota of familial Mediterranean fever disease subjects. *Lett Appl Microbiol* 2017;64(4):254–60. DOI: https://doi.org/10.1111/lam.12722
11. Pepoyan A, Balayan M, Manvelyan A, Galstyan L, Pepoyan S, Petrosyan S et al. Probiotic Lactobacillus acidophilus Strain INMIA 9602 Er 317/402 Administration Reduces the Numbers of Candida albicans and Abundance of Enterobacteria in the Gut Microbiota of Familial Mediterranean Fever Patients. *Front Immunol* 2018;9:1426. DOI: https://doi.org/10.3389/fimmu.2018.01426
12. Hovhannisyan HG, Pashayan MM, Barseghyan AH, Grigoryan GG, Gaboyan EH, Danielyan LV Health promoting potentials of Armenian functional sour milk "Narine" and its starter Lactobacillus helveticus MDC 9602. *NEW Armen Med J* 2023:66–73. DOI: https://doi.org/10.56936/18290825-2023.17.3-66
13. Hovhannisyan HG, Danielyan LV, Baghdasaryan LG, Melkumyan IE, Barseghyan AH. The influence of the alcoholic extract of lemon balm on the growth of Lactobacillus helveticus MDC 9602 and the technological, physicochemical, organoleptic, and antioxidant properties of fermented milk. *J Food Sci Technol* 2025;62(8):1503–12. DOI: https://doi.org/10.1007/s13197-024-06121-x
14. Hovhannisyan HG, Danielyan L, Chichoyan N, Pashayan M, Baghdasaryan L, Melkumyan I et al. Improvement of functional and sensory properties of fermented dairy drink Narine using raw apricot gum. *Funct Foods Health Dis* 2024;14(8):600–14. DOI: https://doi.org/10.31989/ffhd.v14i8.1398
15. Cremonesi. Genome sequence and analysis of Lactobacillus helveticus. *Front Microbiol* 2013. DOI: https://doi.org/10.3389/fmicb.2012.00435
16. Du Y, Xu J, Li J, Wu R. Evaluation of Probiotic Properties and Safety of Lactobacillus helveticus LH10 Derived from Vinegar through Comprehensive Analysis of Genotype and Phenotype. *Microorganisms* 2024;12(4):831. DOI: https://doi.org/10.3390/microorganisms12040831
17. Liu C, Xu L, Chen S, Wang S, Lv M, Dong K et al. Whole-genome sequencing and probiotic properties of Lactobacillus helveticus KM7 isolated from the gut of the Chinese honey bee (Apis cerana): A promising exopolysaccharide-producing strain. *BMC Microbiol* 2025;25(1):530. DOI: https://doi.org/10.1186/s12866-025-04286-9
18. Sereika M, Kirkegaard RH, Karst SM, Michaelsen TY, Sørensen EA, Wollenberg RD et al. Oxford Nanopore R10.4 long-read sequencing enables the generation of near-finished bacterial genomes from pure cultures and metagenomes without short-read or reference polishing. *Nat Methods* 2022;19(7):823–6. DOI: https://doi.org/10.1038/s41592-022-01539-7
19. Thomas C, Brangsch H, Galeone V, Hölzer M, Marz M, Linde J. Accurately assembling nanopore sequencing data of highly pathogenic bacteria. *BMC Genomics* 2025;26(1):783. DOI: https://doi.org/10.1186/s12864-025-11793-6
20. J Shetty P. The Evolution of DNA Extraction Methods. *Am J Biomed Sci Res* 2020;8(1):39–45. DOI: https://doi.org/10.34297/AJBSR.2020.08.001234

21. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 2019;37(5):540–6. DOI: https://doi.org/10.1038/s41587-019-0072-8
22. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB et al. metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat Methods* 2020;17(11):1103–10. DOI: https://doi.org/10.1038/s41592-020-00971-x
23. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30(14):2068–9. DOI: https://doi.org/10.1093/bioinformatics/btu153
24. Shukla SK, Manobala T, Rao TS. The role of S-layer protein (SlpA) in biofilm-formation of Deinococcus radiodurans. *J Appl Microbiol* 2022;133(2):796–807. DOI: https://doi.org/10.1111/jam.15613
25. Jarocki P, Targoński Z. Genetic Diversity of Bile Salt Hydrolases Among Human Intestinal Bifidobacteria. *Curr Microbiol* 2013;67(3):286–92. DOI: https://doi.org/10.1007/s00284-013-0362-1
26. Kim KW, Kang S-S, Woo S-J, Park O-J, Ahn KB, Song K-D et al. Lipoteichoic Acid of Probiotic Lactobacillus plantarum Attenuates Poly I:C-Induced IL-8 Production in Porcine Intestinal Epithelial Cells. *Front Microbiol* 2017;8:1827. DOI: https://doi.org/10.3389/fmicb.2017.01827
27. Kuwana R, Yamamoto N. Increases in GroES and GroEL from Lactobacillus acidophilus L-92 in response to a decrease in medium pH, and changes in cytokine release from splenocytes: Transcriptome and proteome analyses. *J Biosci Bioeng* 2012;114(1):9–16. DOI: https://doi.org/10.1016/j.jbiosc.2012.02.010
28. Cao C, Wang J, Liu Y, Kwok L-Y, Zhang H, Zhang W. Adaptation of Lactobacillus plantarum to Ampicillin Involves Mechanisms That Maintain Protein Homeostasis. *mSystems* 2020;5(1):10.1128/msystems.00853-19. DOI: https://doi.org/10.1128/msystems.00853-19
29. Plaza-Vinuesa L, Sánchez-Arroyo A, Moreno FJ, De Las Rivas B, Muñoz R. Dual 6Pβ-Galactosidase/6Pβ-Glucosidase GH1 Family for Lactose Metabolism in the Probiotic Bacterium *Lactiplantibacillus plantarum* WCFS1. *J Agric Food Chem* 2023;71(28):10693–700. DOI: https://doi.org/10.1021/acs.jafc.3c01158
30. Chen Y, Li P, Huang W, Yang N, Zhang X, Cai K et al. Structural characterization and immunomodulatory activity of an exopolysaccharide isolated from Bifidobacterium adolescentis. *Int J Biol Macromol* 2025;304:140747. DOI: https://doi.org/10.1016/j.ijbiomac.2025.140747
31. D'Aimmo MR, Satti M, Scarafile D, Modesto M, Pascarelli S, Biagini SA et al. Folate-producing bifidobacteria: metabolism, genetics, and relevance. *Microbiome Res Rep* 2023;3(1). DOI: https://doi.org/10.20517/mrr.2023.59
32. Tagliazucchi D, Martini S, Solieri L. Bioprospecting for Bioactive Peptide Production by Lactic Acid Bacteria Isolated from Fermented Dairy Food. *Fermentation* 2019;5(4):96. DOI: https://doi.org/10.3390/fermentation5040096
33. Jung G, Zin H, Son B, Shin H, Kim J. Characterization of a plasmid dependent DNA phage targeting Escherichia coli harboring a conjugative plasmid and its impact on gut microbiota. *Sci Rep* 2025;15(1):19701. DOI: https://doi.org/10.1038/s41598-025-04494-3
34. Ku H, Kelk D, Bauer DC, Sidhu JPS. Phage-plasmid hybrids as vectors for antibiotic resistance in environmental Escherichia coli. *Sci Total Environ* 2025;959:178157. DOI: https://doi.org/10.1016/j.scitotenv.2024.178157
35. Forsyth JH, Barron NL, Scott L, Watson BNJ, Chisnall MAW, Meaden S et al. Decolonizing drug-resistant E. coli with phage and probiotics: breaking the frequency-dependent dominance of residents. *Microbiology* 2023;169(7). DOI: https://doi.org/10.1099/mic.0.001352