Comparative Genomic Analysis of *Limosilactobacillus Pontis* Reveals Genetic Characteristics of High GC Content and Potential Niches Adaptations

Qing Liu (✉ liuqing0471@163.com)

Inner Mongolia Agricultural University

Chen Liu

Inner Mongolia Agriculture University: Inner Mongolia Agricultural University

Weicheng Li

Inner Mongolia Agriculture University: Inner Mongolia Agricultural University

Wenjun Liu

Inner Mongolia Agriculture University: Inner Mongolia Agricultural University

Qing Liu

Inner Mongolia Agriculture University: Inner Mongolia Agricultural University <https://orcid.org/0000-0001-7970-5887>

Research Article

Keywords: Limosilactobacillus pontis, Comparative genomic, Niches adaptability

Posted Date: October 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-969439/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Limosilactobacillus pontis is a species of lactic acid bacteria (LAB) found in fermented milk, sourdough and broiler chickens gastrointestinal tract and so on. However, the evolutionary strategies and genomic characteristics of the species remain unknown, which limits its application. In this study, whole genome sequencing was combined with a comparative genomic approach to investigate genomic characteristics and evolutionary strategies of *L. pontis*; this includes three published genomic sequences and two strains isolated from fermented milk in Inner Mongolia, China. The mean genome size and GC content of *L. pontis* was 1.70 Mb and 53.06%, respectively. Within the LAB *L. pontis* has a high GC content. The phylogenetic tree based on 1,281 core genomes showed that strains from the same sources aggregated together in clusters. Genome information, average nucleotide identity values, and phylogenetic relationships amongst *L. pontis* from different sources indicated that strains have potential niche adaptability. Functional genomic aspects, GT2 and GT4 (glycosyltransferases, GTs) involved in the synthesis of cellulose and sucrose were the family with the largest number of carbohydrate enzymes in *L. pontis*, particularly strains isolated from fermented milk. It is worth mentioning that the ability of *L. pontis* to produce bacteriocin may increase its application potential. This study provides new insight into the genetic characteristics and potential niche adaptations of *L. pontis*.

Highlight

1. In this study, whole genome sequencing was combined with a comparative genomic approach to investigate genomic characteristics and evolutionary strategies of *L. pontis*.
2. The ability of *L. pontis* to produce bacteriocin may increase its application potential.
3. The GC content of *L. pontis* was higher than a lot of LAB. In bacteria a high GC content is related to a faster growth rate and a more stable DNA structure.
4. Comparative genomic analysis of *Limosilactobacillus pontis* reveals potential niches adaptations

Introduction

Limosilactobacillus pontis (*L. pontis*) is commonly found in fermenting animal and plant materials, such as sourdough, fermented milk, vinegar and the gastrointestinal tract of broiler chickens (Vogel et al. 1994). The species, *L. pontis*, previously belonged to the genus *Lactobacillus*. Due to the extreme phenotypic, ecological and genotypic diversity amongst species within the genus *Lactobacillus*, the genus was recently reclassified into 25 genera, including *Lactobacillus* (previously the *L. delbrueckii* group and *Limosilactobacillus* (Zheng J 2020). *Limosilactobacillus* includes *Limosilactobacillus reuteri* and *Limosilactobacillus fermentum* amongst others. Being 'generally recognized as safe' (Maya-Barrios et al. 2021), strains in this species have been widely applied in various food fermentations (Bao et al. 2010). Furthermore, *L. reuteri* is a special member of the LAB (Saviano et al. 2021) as it can convert glycerol into a broad-spectrum antibacterial compound, called reuterin (Talarico et al.); secretion conditions include an appropriate temperature, pH and anaerobic environment. It has been shown experimentally that reuterin is

a mixture of monomers, hydrated monomers and cyclic dimers of L-hydroxypropanal (Arqués et al. 2004). It has a large antibacterial range and can inhibit the growth of Gram-positive bacteria, Gram-negative bacteria and fungi (Dobrogosz and Lindgren 1995). Because *L. pontis* belongs to the same genus as *L. reuteri*, it may also have potential bacteriocin-producing characteristics. Despite being reported from a wide range of sources, there have been few studies on *L. pontis* and it is worthy of further exploration.

Comparative genomic studies of strains within the same species have provided insight into genetic features that have been modified, acquired, or lost during evolution and adaptation to specific environmental niches (Feyereisen et al. 2019). In recent years, along with the rapid development of high-throughput DNA sequencing technologies, genomic studies have become more feasible and affordable; as a result genomic data from numerous organisms has become publicly available (Douillard et al. 2013). To find environmentally relevant features of the lactic acid genome, Mendes-Soares, et al., used comparative genomic analysis to examine 67 strains from food and the intestines, and reproductive tract of humans. The results showed that the genome and the content of GC of LAB in the human reproductive tract was significantly smaller and lower than other strains (Denou et al. 2008). *Limosilactobacillus reuteri* can be isolated from various sources including the intestines of pigs, humans and rodents. In an investigation of *L. reuteri* from different sources, the phylogenetic tree based on core-genome features had different evolutionary branches with habitats as clusters. This indicated that *L. reuteri* from different genetic backgrounds could selectively colonize and adapt to specific habitats; this ability was significantly associated with acid and bile salt tolerance, and the cell adhesion ability of *L. reuteri* (Oh et al. 2010). The genetic and evolutionary relationships amongst strains of *Lactocaseibacillus rhamnosus* have also been evaluated and showed that genes involved in glucose metabolism in *L. rhamnosus* were significantly correlated with the isolation source (Douillard et al. 2013). *Ligilactobacillus salivarius* also showed close phylogenetic relationships with habitat, including a specific connection related to its ability to produce bacteriocin (Lee et al. 2017). Identification of bacteriocin *in vitro* and *in vivo* is a time-consuming and cumbersome procedure; in contrast, *in silico* exploration of genes encoding bacteriocin biosynthesis provides a rapid and effective way to identify putative bacteriocin (Azevedo et al. 2015). However, current understanding of *L. pontis* genomic information is poor, limiting its application; comparative genomics could be used to reveal its genomic characteristics in order to gain a better understanding of the strain.

In this study we obtained two strains of *L. pontis* from Lactic Acid Bacteria Collection Center (LABCC) that had been derived from fermented milk: IMAU10341 and IMAU10345. We determined the genome sequence of both strains using Illumina Novaseq 6000 genome sequencing. We downloaded from NCBI (<http://www.ncbi.nlm.nih.gov>) the genome sequences of the type strain of *L. pontis* DSM8475^T and three other strains: *L. pontis* LP475, *L. pontis* UMB0683, *L. pontis* MAG.251. We used a comparative genomic approach to define the pan-genome, core-genome, and unique genes of strains from sourdough, fermented milk and broiler chicken gastrointestinal tract, assessed genetic diversity and gained insight into the distinguishing features of particular isolates, compared the numbers of functional genes and carbohydrate-active enzymes (CAZymes), and explored expression of the gene responsible for bacteriocin

production. The aim of the current work was to carry out comparative genomics analyses of *L. pontis* isolated from different niches, and to explore the genetic diversity and potential host adaptation of the species.

Materials And Methods

1. Strains

The two strains evaluated in this study were from the Lactic Acid Bacteria Collection Center (LABCC) of the Inner Mongolia Agricultural University; IMU10341 and IMU10345 were isolated from fermented milk produced in Tumao Ermu, East Wukuriye, Qinqi Qi, East Wuzumu, Xilingol League, Inner Mongolia (Liu et al. 2012). The genomes of a further nine strains for comparison were downloaded from the NCBI Refseq (National Coalition Building Institute, <https://www.ncbi.nlm.nih.gov/refseq/>) database.

2. Main instruments and reagents

Genomic DNA extraction kit, Beijing Tiangen Biochemical Technology Co, Ltd; MRS medium(lactic acid bacteria culture medium), Oxoid Company. An electric constant temperature incubator, Beijing Yiheng Technology Co, Ltd. An ultra-micro UV spectrophotometer, NanoDrop company (Jie et al. 2018).

3. Strain culture and genomic DNA extraction

An appropriate amount of the freeze-dried bacterial powder was inoculated into 5 mL of MRS liquid medium and incubated at 37 °C for 24 h. Then inoculate 2% of the inoculum was inoculated into either 5 mL of liquid MRS medium, 15 mL of MRS liquid medium and incubated for a further 24 hours at 37 °C in an anaerobic workstation, after which the bacteria were collected by centrifugation (Jie et al. 2018). A genomic DNA extraction kit was used to extract genomic DNA following the manufacturer's instructions. The purity and quality of the extracted DNA was checked using an ultra-micro ultraviolet spectrophotometer (Tanaka et al. 2009).

4. Genome sequencing and assembly

If the concentration and purity met sequencing requirements, an Illumina Novaseq 6000 sequencing platform was used to complete the whole genome sequencing, and the library constructed by referring to the literature; each strain was constructed separately using a PE150 library. The quality of the raw data obtained was evaluated. To obtain high quality reads the low-quality reads were filtered out and adapters and primers removed. The high-quality reads were spliced using the software package SOAPdenovo V2.0 (Luo et al. 2012). This software was also used for filling and single base correction, to achieve complete assembly of the genome (Xu et al. 2020).

5. Comparative Genomic Analysis

5.1 Average nucleotide identity (ANI) and total nucleic acid identity (TNI)

In this study, ANI and TNI of *L. pontis* were calculated using the method of Chen et al. (Chen et al. 2015). A self-made Perl script was used for calculation.

5.2 Construction of core-genome set and accessory gene set analysis

After gene prediction of the strain genome based on Prokka (Seemann 2014) software, Roary (Page et al. 2015) software was used to identify the core-genome set and accessory-gene set. The core-genome was identified based on the principle that the amino acid similarity of the encoded protein was greater than 95%. The core-genome (Duranti et al. 2016) is comprised of the genes that exist in the genome of all strains. Any gene that appears in only one strain is an unique gene.

5.3 Phylogenetic tree construction

We used the core-genome sequence obtained by Roary to construct a phylogenetic tree on the ITOL online website (<https://itol.embl.de/>) to explore phylogenetic relationships amongst *L. pontis* isolates (ca et al. 2019).

5.4 Functional gene annotation

All genome sequences of the *L. reuteri* strains were annotated by uploading each sequence file to the Rapid Annotation using Subsystem Technology (RAST) web server (Aziz et al. 2008) for genomic annotation. The protein coding sequences were predicted and categorized based on the SEED subsystem (most prominently FIGfams) in RAST.

5.5 CAZyme

Carbohydrate active enzymes were annotated and compared for the genome sequences of all strains using the CAZyme database. According to the results, the detailed carbohydrate active enzyme family was searched on the CAZy (<http://www.cazy.org/>) webpage information (Cantarel 2009).

5.6 Bacteriocin

BAGEL4 is an online database that helps mine and visualize ribosome-synthesized and post-translationally modified peptides and bacteriocin-producing gene clusters in the prokaryotic genome (<http://bagel4.molgenrug.nl/index.php>). On the BAGEL4 web server, a DNA nucleotide sequence was used as an input file (J et al. 2018).

6. Uploading the genomes of strains

The assembled genomic data of the two strains of *L. pontis* (IMAU10341 and IMAU10345) were uploaded to the SRA database in NCBI. The Whole Genome Shotgun sequences of strains IMAU10341 and IMAU10345 project has been deposited at DDBJ/EMBL/GenBank under the accession JAIUWJ000000000 and JAIUWK000000000, respectively.

Results

General Genomic Characteristics

The genome size of the five strains of *L. pontis* ranged from 1.60-1.77 Mb; the average size was 1.70 Mb, and the average (G+C) mol% content was 53.06%. The size of the assembled genome of *L. pontis* strain IMAU10341 was 1.77 Mb, encoding 1787 genes, and the (G+C)mol% content was 52.49%; the size of the assembled genome of *L. pontis* strain IMAU10345 was 1.76 Mb, encoding 1786 genes, and the (G+C) mol% content was 52.49%. In addition, the number of tRNAs in each genome of *L. pontis* was predicted, and the average number of tRNAs was 51. The (G+C) mol% content of *L. pontis* UMB0683 was 48.70%, and the genome size was 1.87 Mb, which did not match the characteristics of the *L. pontis* genome and requires further analysis.

ANI and TNI

The ANI is a classical method for identifying unique species or potential subspecies existing amongst strains within the same species. ANI values of the six genomes were made through pairwise comparison at the 95% threshold to further identify the species to which they belonged ([Goris et al. 2007](#)).

In this study, ANI and TNI were analyzed on the whole genome sequence of six *L. pontis* strains, and a heat map was drawn (Fig. 1). The results showed that, *L. pontis* UMB0683 the value of ANI was 83.99% less than 96.00%, the value of TNI was 67.41% less than 78.00%. Therefore, *L. pontis* UMB0683 may not be a strain from this genus, and was excluded from further analysis. The remaining four strains of *L. pontis* and the type strain *L. pontis* DSM8475^T had ANI values greater than 97.00% and TNI values greater than 78.00%, and were used for subsequent analysis. The ANI between IMAU10341 and IMAU10345 was 99.98%, and the TNI was 99.57%. The ANI between IMAU10341 and DSM8475^T was 98.88%, the ANI between IMAU10345 and DSM8475^T was 98.97%, the ANI between MAG.251 and the type strain was 97.79%, and the ANI between LP475 and DSM8475^T isolated from the same type strain was 99.95%; the degree of matching of base and nucleic acid was high.

Core-Genome and Pan-Genome Analysis

In order to compare the genetic differences between strains, analysis was done using Prokka and Roary software, with greater than 90% amino acid identity as the standard, and the number of core-genome and pan-genes was counted. The pan-gene set of five strains of *L. pontis* contained 2166 genes, of which 1,281 were core-genome. It can be seen from the Venn ([Johnson 1991](#)) diagram that there were: eight and seven unique genes derived from fermented milk strains IMAU10341 and IMAU10345 which were all pseudogenes; 37 and 52 unique genes in sourdough-derived strains DSM8475^T and LP475; and 294 specific genes in animal-derived strain MAG.251.

Phylogenetic tree

Phylogenetic trees directly show the genetic distances and evolutionary relationships amongst individuals in the same population. Building a phylogenetic tree is a necessary prerequisite for studying population structure and species evolution [41]. To further understand the population structure of *L. pontis*, we used Neighbor-Joining to construct a phylogenetic tree, with a bootstrap value of

1000 (Balakirev et al. 2012). This was based on a comparison of the nucleic acid sequence of 1,281 core-genome of *L. pontis* with closely-related species as outgroups: *L. antri* DSM16041^T, *L. frumenti* DSM13145^T, *L. oris* DSM4864^T, *L. panis* DSM6035^T, *L. reuteri* DSM20016^T, *L. vaginalis* DSM5837^T, *L. equigenersi* DSM18793^T, *L. gastricus* DSM16045^T, *L. fermentum* DSM20052^T, *L. coleohominis* DSM14060^T, *L. ingluviei* DSM15946^T, *L. mucosae* DSM13345^T. The five strains of *L. pontis* were very different to each due to the development of genetic diversity during the evolution process; they were divided into three main branches (Fig. 3). The branch where *L. pontis* MAG.251 was located was closer to the root of the outer group, and the other five strains. The bacteria are divided into three branches: strains from fermented milk (IMAU10341 and IMAU10345) were in one branch, and isolates from sourdough (LP475 and DSM8475^T) were together in another branch.

Functional gene analysis

RAST

The *L. pontis* strains evaluated had more genes related to metabolism of nucleosides and nucleotides, protein, amino acids and derivatives, and carbohydrates than other type of genes (Fig. 4). Isolates *L. pontis* IMAU10341 and *L. pontis* IMAU10345 from fermented milk, had significantly more genes involved in metabolism of carbohydrates, amino acids and derivatives, fatty acids, lipids, and isoprenoids than the other strains; genes involved in cell wall and capsule development, toxicity, disease and defense were also more abundant in these two strains than the other isolates. Genes related to protein metabolism were highest in the sourdough isolate, LP475.

CAZY

In order to further evaluate differences in the genomes of *L. pontis* strains and explore the reasons for *L. pontis* diversity, we used prokka annotation and RAST annotation together. After their combined use, we found that the proportion of carbohydrate utilization related functions in each genome was significant and the number of related functional genes was quite different amongst strains. The abundance of carbohydrate-related active enzymes in the strains was compared. The Carbohydrate Active Enzyme Database (CAZy) is an enzyme database resource, based on the similarity of amino acid sequences in protein domains, which can synthesize or decompose various complex carbohydrates and sugar complexes (Accetto and Avcu?Tin 2015). The main function of the GH family is to hydrolyze or rearrange glycosidic bonds and includes 152 gene families; the GT family is mainly responsible for formation of glycosidic bonds and includes 105 gene families; the function of the PL family is to break glycosidic bonds (not by hydrolysis) and includes 28 gene families; the CE family can hydrolyze carbohydrate esterases and includes 16 gene families; the AA family is an oxidoreductase, which works in synergy with carbohydrate active enzymes and includes 14 gene families; the CBM family is attached to carbohydrates and includes 83 gene families (Cao et al.). Predictions from the five strains of *L. pontis* included 11 GH, eight GT and one CE, as well as the distribution and abundance of GH, GT and CE genes in all *L. pontis* strains (Fig. 5).

Bacteriocin

In recent years, the use of software combined with databases can quickly identify operons that can encode bacteriocins in the bacterial genome. Most studies use BAGEL to identify potential bacteriocin operons. This study also used BAGEL and predicted class I and class II bacteriocins in *L. pontis* IMAU10341, IMAU10345, and DSM8475^T (Table.2). Strain MAG.251 contained a class II bacteriocin. Class II bacteriocins are small thermostable peptides, which are further subdivided into IIa, IIb, IIc, and IId based on the structure and activity of the peptide. IIa bacteriocins show a narrow range of antimicrobial activity. Class IIb is composed of two unmodified peptide bacteriocins and its activity depends on the synergistic activity of the two peptides. Class IIc bacteriocins are also called cyclic bacteriocins (Klaenhammer 1993). IMAU10341, IMAU10345 and DSM8475^T produced nisin which is a polypeptide that, together with acids and peroxides, is among the biological inhibitors produced by *Lactococcus lactis*. Fermented milk is rich in nutrients and contains many types and number of strains, so the strains need to produce bacteriocins for niche competition.

Discussion

In this study, comparative genomics analysis was done on the whole genome sequence of five strains of *L. pontis* from different sources. It was found that *L. pontis* has a smaller genome and a higher G + C% content than a lot of LAB. The length of genome sequence of different LAB is between 1.80 and 2.90 Mb, and the genome size of *L. pontis* is within that range. The GC content of most LAB genomes is below 50%, the GC content of *Ligilactobacillus salivarius* UCC118 (Netherlands 2008) is at least 33%. The GC content of *L. pontis* was higher than a lot of LAB. In bacteria a high GC content is related to a faster growth rate and a more stable DNA structure (Bobay and Ochman 2017); not only is the genome density relatively large, but also the ability to resist high temperature and alkaline environment is enhanced (Musto et al. 2004). Yue Xiao (Xiao et al. 2021) observed a reduction in the GC content in host-adapted *Lactobacillus* species compared with nomadic species and free-living species, this is contrary to the results of this study.

Based on the core-genome phylogenetic tree, it was found that isolates from fermented milk, sourdough and chicken gastrointestinal tract were in different branches; the branch that included isolates from chicken gastrointestinal tract was closer to the common ancestor. Conclusions from ANI and TNI analysis were consistent with those from phylogenetic tree analysis. The ANI values of IMAU10341 and IMAU10345 were high and different to the ANI values of the type strains. The two strains isolated from similar samples were closely related, but not identical. Strains from fermented milk, sourdough and broiler chicken gastrointestinal tract naturally clustered into several categories. In addition, there were some differences in ANI values between strains from different sources and the type strain, and the ANI values of *L. pontis* strains from different hosts indicate that the species have habitat adaptability.

In this study, the number of genes specific to a strain were similar when strains were from the same source, but differed greatly when strains were from different sources. The unique genes of *L. pontis* were

under positive selection to adapt to different environments, and this study demonstrated that food and intestinal habitats exerted different selective pressures that were related to growth rate and metabolism (Maldonado-Gómez et al. 2016). The unique genes of IMAU10341 and IMAU10345 were pseudogenes. Some original genes can be degraded or inactivated if they are no longer needed. The formation of pseudogenes is a mechanism for gene decline or deletion in the genome (Lerat 2005). We speculate that fermented milk is nutritious and suitable environment, which leads to the phenomenon of genome degradation in strains compared with strains from other environments.

Shuo Wang, et al., (Wang et al. 2020) evaluated the phylogenetic relationships amongst *L. ruminis* strains and created a phylogenetic tree based on orthologous genes from 91 genomes that constituted the core-genome. They found that strains from porcine, bovine and human faeces were divided into three clades; from the phylogenetic tree strains from the same sample type were clustered, closely related, and with the potential for niche adaptation. In our study, MAG.251 was closest to the root, indicating that its genome is closer to its ancestors than the other strains evaluated. In different environments, the evolutionary process of species adaptation is different. Adaptation of the bacterial genome to specific environments will accumulate genetic variation related to those environmental characteristics. The genomes of the other four strains all changed due to environmental changes. Strains from different sources clustered together, which indicated that the phylogenetic relationship reflected potential niches adaptation which was consistent with ANI results. Adaptive evolution of microorganisms to specific niches is also reflected in their functional genomics, and specific niches will shape specific functional genomics. In order to reduce reproduction-related energy consumption, strains living in specific niches will lose redundant functional genes, and at the same time evolve specific genes adapted to specific niches (Jung et al. 2018). Functional gene annotation results found that *L. pontis* isolates from fermented milk (IMAU10341 and IMAU10345) were rich in carbohydrate-related genes and genes involved in amino acids and derivatives. Genes related to metabolism of fatty acids, lipids and isoprenoids were significantly more abundant in fermented milk isolates than in other strains, while genes for protein metabolism were more abundant in the sourdough isolate LP475 than in other strains.

Based on RAST annotations, carbohydrates are the main energy source for growth and development of organisms. Studies have shown that LAB can specifically adapt to the environment by obtaining carbohydrate-related genes from the environment or losing non-essential carbohydrate genes. Differences in carbohydrate genes amongst the five *L. pontis* strains may be related to their different isolation sources; strains from the same isolation source were very similar. Those isolated from fermented milk had more cell wall and capsule genes than other strains. This may be due to the rich environment of fermented milk that also contains many other microorganisms; increasing the thickness of the cell wall would help resist external antibacterial substances.

LAB require exogenous amino acids and peptides which are provided by hydrolysis of casein in fermented milk. Many LAB are deficient in a variety of amino acids (Walter et al. 2011). In addition to enabling microbial growth, peptides, amino acids and their derivatives contribute to the formation of the

unique texture and flavour of fermented milk, so strains from fermented milk contain more genes associated with metabolism of amino acids and their derivatives.

In addition, all *L. pontis* strains contained genes for producing extracellular polysaccharides, which protect them against external stresses. Because *L. pontis* contains more genes associated with carbohydrate metabolism, we also undertook carbohydrate enzyme analysis. CAZY annotation showed that the distribution and abundance of CE, GH, and GT families of genes in *L. pontis* were visualized by heat map and clustered using a hierarchical clustering method. This found that *L. pontis* was divided into three groups, representing the three different isolation sources. There are differences in the composition of carbohydrate active enzymes. The strains of *L. pontis* contained eight GT families, amongst which GT2 and GT4 were the most widely distributed. GT2 is a diverse family. Either the sugar is changed from UDP-glucose, UDP-N-acetyl-galactosamine, GDP-mannose, or CDP-varying sugars are transferred to a range of substrates including cellulose, dihydroethanol phosphate and phosphoric acid. Studies have shown that GT2 and GT4 are rich in glycosyltransferases in naturally fermented mares' milk. In fermented milk, these substances, especially polysaccharide compounds, may be related to the final viscosity and water holding capacity of fermented milk products. The *L. pontis* strains were rich in GT2 and GT4, indicating that they may have higher sugar synthesis ability. In addition, the GT32 family, which participates in conversion of mannose, glucose and galactose, and the GT101 family, which represents β -glucosyltransferase, were only present in MAG.251; this may be related to its animal origin. GH2 mainly encodes β -glycosidases, these enzymes are usually related to metabolism of lactose and galactose, GH13 mainly encodes α -glycosidases, which are related to the metabolism of starch, and GH25 and GH73 mainly encode lysozymes, which are related to antibacterial activities of bacteria (originally found in fungi or bacteriophages). This kind of family has a large number of strains from herbivores, kefir and porridge. The possible reasons for this is that the genetic changes made by the strains to resist external environmental factors indicate that they can resist external change. GH53 was only found in sourdough isolates, which may be related to its ability to metabolise hemicellulose (Douillard et al. 2013).

Bacteriocins are bacteriostatic substances produced by LAB in order to compete for ecological niches. Amongst the five strains of *L. pontis*, four that were derived from fermented milk, sourdough and broiler chicken gastrointestinal tract, produced a lot of bacteriocins. IMAU10341, IMAU10345 and DSM8475^T produced nisin which is a polypeptide that, together with acids and peroxides, is among the biological inhibitors produced by *Lactococcus lactis*. One potential function of nisin is that it has a regulatory action on the growth cycle of the producing organism (Ruyter et al. 1996), enabling them to compete under real-life conditions (Wang et al. 2020). Results from functional genome analysis were the same as the phylogeny results, and gene expression of the different strains was similar, reflecting potential niche adaptation. At present, there are few published genomes for *L. pontis* which limits in-depth study of the species, and it may be necessary to further explore appropriate culture conditions, increase the number of strains isolated from various sources and samples for research.

The results of this study provide a reference for understanding the evolution of *L. pontis*. Phylogenetic and functional genomic characteristics prove that *L. pontis* strains have potential niche adaptations.

However, due to the limited number of strains available, further comparative genomic research is needed for verification.

Conclusions

In this study, the genome sequences of five *L. pontis* strains provided a basis for analysis of this species. *L. pontis* is characterized by high GC content and potential niches adaptations which is different to the characteristics presented in the literature previously. Moreover, the strains produced different functional genes which were related to their different isolation sources, such as bacteriocin gene, which enables adaptation to various environments. In summary, *L. pontis* is a high GC content strain of LAB and has potential niche adaptations.

Declarations

Conflict of interest

We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Author contributions

Qing Liu - conceptualization, investigation, visualization, writing - original draft, review & editing; Chen Liu - conceptualization, investigation, writing - original draft, review & editing; Weicheng Li - investigation (physiological, biochemical analyses and chemotaxonomy); Wenjun Liu - suggestions, review and editing; Heping Zhang - resources and supervision.

References

1. Tomaž A, Gorazd A (2015) Polysaccharide utilization locus and CAZYme genome repertoires reveal diverse ecological adaptation of *Prevotella* species. *Systematic Applied Microbiology* 7:453–524
2. Juan LA, Fernández J, Gaya P, Nuñez M, Rodríguez E, Medina M (2004) Antimicrobial activity of reuterin in combination with nisin against food-borne pathogens. *Int J Food Microbiol* 95:225–229
3. Azevedo AC, Bento C, Ruiz JC, Queiroz MV, Mantovan HC (2015) Distribution and Genetic Diversity of Bacteriocin Gene Clusters in Rumen Microbial Genomes. *Appl Environ Microbiol* 20:290–304
4. Aziz RK et al (2008) The RAST Server: Rapid Annotations using Subsystems Technology. *Bmc Genomics* 9:75–75
5. Balakirev ES, Krupnova TN, Ayala FJ (2012) Concatenate_NJ_tree_Saccharina_rbcLS_COI_ITSDData from: DNA variation in the phenotypically-diverse brown alga *Saccharina japonica*. *Encyclopedia of Genetics, Genomics, Proteomics and Informatics* 1:108–108
6. Bao Y et al (2010) Screening of potential probiotic properties of *Lactobacillus fermentum* isolated from traditional dairy products. *Food Control* 21:695–701

7. Bobay LM, Ochman H (2017) Impact of Recombination on the Base Composition of Bacteria and Archaea. *Mol Biol Evol* 34:2627–2636
8. Ca I, Letunic, Peer B (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic acids research* 47:256–259
9. Cantarel B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nuc Acids Res* 37:233–238
10. Cao H, Ekstrom A, Yin Y Plant Carbohydrate Active Enzyme (CAZyme) Repertoires: A Comparative Study. Springer International Publishing 8:115–134
11. Chen J et al (2015) SISP: a Fast Species Identification System for Prokaryotes Based on Total Nucleotide Identity of Whole Genome Sequences. *Infectious Diseases Translational Medicine* 1:8–8
12. Denou E, Pridmore RD, Berger B, Panoff JM, rigoni FA, Brussow H (2008) Identification of genes associated with the long-gut-persistence phenotype of the probiotic *Lactobacillus johnsonii* strain NCC533 using a combination of genomics and transcriptome analysis. *J Bacteriol* 190:3161–3168
13. Dobrogosz WJ, Lindgren SE (1995) Antibiotic reuterin. WO 1988008452:A1
14. Douillard FP et al (2013) Comparative Genomic and Functional Analysis of 100 *Lactobacillus rhamnosus* Strains and Their Comparison with Strain GG. *Plos Genetics* 9:e1003683
15. Duranti S et al (2016) Evaluation of genetic diversity among strains of the human gut commensal *Bifidobacterium adolescentis*. *Sci Rep* 6:23971
16. Feyereisen M et al (2019) Comparative genome analysis of the *Lactobacillus brevis* species. *BMC Genom* 20:1–15
17. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Tiedje JM (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 57:81–91
18. van HA J, Anne de J, Song C, Viel JH, Jan K, Kuipers OP (2018) BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nuclc Acids Research* W1
19. Jie Y et al (2018) Comparative Genomics of the Herbivore Gut Symbiont *Lactobacillus reuteri* Reveals Genetic Diversity and Lifestyle Adaptation. *Front Microbiol* 9:1151–1151
20. Johnson B (1991) Tree-maps: a space-filling approach to the visualization of hierarchical information structures. In: *Proceeding Visualization '91*
21. Jung J et al (2018) Comparative genomic analysis of *Lactobacillus plantarum* GB-LP4 and identification of evolutionarily divergent genes in high-osmolarity environment. *Genes Genomics* 40:217–223
22. Klaenhammer TR (1993) Genetics of bacteriocin produced by LAB. *FEMS Microbiol Rev* 12:39–85
23. Lee JY, Han GG, Kim EB, Choi YJ (2017) Comparative Genomics of *Lactobacillus salivarius* Strains Focusing on Their Host Adaptation. *Microbiol Res* 205:48–58
24. Lerat (2005) Recognizing the pseudogenes in bacterial genomes. *Nucleic Acids Res* 33:3125–3132

25. Liu W et al (2012) Isolation and identification of lactic acid bacteria from Tarag in Eastern Inner Mongolia of China by 16S rRNA sequences and DGGE analysis. *Microbiol Res* 167:110–115
26. Luo R, Liu B, Xie Y, Li Z, Liu Y (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* 1:18–18
27. Maldonado-Gómez M et al (2016) Stable Engraftment of *Bifidobacterium longum* AH1206 in the Human Gut Depends on Individualized Features of the Resident Microbiome. *Cell Host Microbe* 4:515–526
28. Maya-Barrios A, Lira-Hernandez K, Jiménez-Escobar I, Hernández L, Gutiérrez-Castrellón P (2021) *Limosilactobacillus reuteri* ATCC PTA 5289 and DSM 17938 as adjuvants to improve evolution of pharyngitis/tonsillitis in children: randomised controlled trial. *Beneficial Microbes* 12:1–10
29. Musto H, Naya H, Zavala A, Romero H, Alvarez-Valín F, Bernardi G (2004) Correlations between genomic GC levels and optimal growth temperatures in prokaryotes. *FEBS Letters* 573:73–77
30. Netherlands S (2008) *Lactobacillus salivarius* UCC118. Springer Netherlands 2018:1076–1076
31. Oh PL et al (2010) Diversification of the gut symbiont *Lactobacillus reuteri* as a result of host-driven evolution. *ISME Journal* 4:377–387
32. Page AJ et al (2015) Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 22:3691–3693
33. Ruyter D, Pascalle G, Kuipers, Oscar P (1996) Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin. *Applied Environmental Microbiology* 10:3662–3667
34. Saviano A, Brigida M, Migneco A, Gunawardena G, Ojetti V (2021) *Lactobacillus Reuteri* DSM 17938 (*Limosilactobacillus reuteri*) in Diarrhea and Constipation: Two Sides of the Same Coin? *Medicina* (Kaunas, Lithuania) 57:643–643
35. Seemann T (2014) Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* 14:2068–2069
36. Talarico TL, Casas IA, Chung TC, Dobrogosz WJ Production and isolation of reuterin, a growth inhibitor produced by *Lactobacillus reuteri*. *Antimicrobial Agents and Chemotherapy* 12:1854–1858
37. Tanaka T, Sakai R, Kobayashi R, Hatakeyama K, Matsunaga T (2009) Contributions of Phosphate to DNA Adsorption/Desorption Behaviors on Aminosilane-Modified Magnetic Nanoparticles. *Langmuir* 25:2956–2961
38. Vogel RF et al (1994) Identification of *Lactobacilli* from sourdough and description of *Lactobacillus pontis* sp. nov. *Int J Syst Bacteriol* 44:223–229
39. Walter J, Britton RA, Roos S (2011) Host-microbial symbiosis in the vertebrate gastrointestinal tract and the *Lactobacillus reuteri* paradigm. *Proc Natl Acad Sci USA* 1:4645–4652
40. Wang S, Yang B, Ross RP, Stanton C, Chen W (2020) Comparative Genomics Analysis of *Lactobacillus ruminis* from Different Niches. *Genes* 11:70–70
41. Xiao Y, Zhao J, Zhang H, Zhai Q, Chen W (2021) Mining genome traits that determine the different gut colonization potential of *Lactobacillus* and *Bifidobacterium* species. *Microbial Genomics* 6:1–17

42. Xu M, Guo L, Gu S, Wang O, Zhang Y (2020) TGS-GapCloser: A fast and accurate gap closer for large genomes with low coverage of error-prone long reads. *GigaScience* 9:1–11
43. Zheng JWS, Salvetti E, Franz CMAP, Harris HMB et al (2020) A taxonomic note on the genus *Lactobacillus*: Description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. *Int J Syst Evol Microbiol* 70:2782–2858

Tables

Table 1
List of the *L. pontis* strains evaluated in this study

Strain	Isolation source	CDSs	Size (Mb)	(G + C)mol(%)	rRna	tRna	Accession number
<i>L. pontis</i> IMAU10341	fermented milk	1787	1.77	52.49	6	64	JAIUWJ000000000
<i>L. pontis</i> IMAU10345	fermented milk	1786	1.76	52.49	6	64	JAIUWK000000000
<i>L. pontis</i> DSM8475 ^T	sourdough	1539	1.67	53.40	3	29	GCA_001435345.1
<i>L. pontis</i> LP475	rye sourdough	1561	1.71	53.30	18	67	GCA_009428965.1
<i>L. pontis</i> UMB0683	catheter	1698	1.87	48.70	7	59	GCA_002940945.1
<i>L. pontis</i> MAG.251	broiler chicken gastrointestinal tract	1551	1.60	53.60	1	32	GCA_017888165.1
<i>L. antri</i> DSM16041 ^T	stomach mucosa	2038	2.30	51.00	2	45	GCA_000160835.1
<i>L. frumenti</i> DSM13145 ^T	rye-bran sourdough	1633	1.73	42.60	3	40	GCA_001436045.1
<i>L. oris</i> DSM4864 ^T	saliva	1861	2.03	50.00	2	51	GCA_001434465.1
<i>L. panis</i> DSM6035 ^T	sourdough	1797	2.01	48.10	2	33	GCA_001435935.1
<i>L. reuteri</i> DSM20016 ^T	intestine of adult	1889	2.00	38.90	18	68	GCA_000016825.1
<i>L. vaginalis</i> DSM5837 ^T	vaginal swab	1681	1.79	40.50	3	24	GCA_001435915.1
<i>L. equigenerosi</i> DSM18793 ^T	thoroughbred horses	1528	1.60	42.70	1	41	GCA_001435245.1
<i>L. gastricus</i> DSM16045 ^T	gastric biopsies, stomach mucosa	1761	1.85	41.60	3	37	GCA_001434365.1
<i>L. fermentum</i> DSM20052 ^T	fermented beets	1719	1.89	52.50	15	59	GCA_013394085.1

Strain	Isolation source	CDSs	Size (Mb)	(G + C)mol(%)	rRna	tRna	Accession number
<i>L. coleohominis</i> DSM14060 ^T	vagina	1715	1.72	40.80	3	29	GCA_001435055.1
<i>L. ingluviei</i> DSM15946 ^T	crop	2055	2.15	49.90	2	69	GCA_001435775.1
<i>L. mucosae</i> DSM13345 ^T	small intestine	1990	2.25	46.40	1	42	GCA_001436025.1

Table 2
Bacteriocins in the five *L. pontis* strains

Strain	Bacteriocin
IMAU10341	1.Nisin_U 2.rSAM-modified_RiPP_008 3.Divercin_V41_(DvnV41) 4.Divercin_RV41(mature)
IMAU10345	1.SSV-2083 2.Nisin_U 3.rSAM-modified_RiPP_078 4.rSAM-modified_RiPP_040 5.rSAM-modified_RiPP_041 6.rSAM-modified_RiPP_089
DSM8475 ^T	1.Nisin_U 2.Pinensin_10 3.Blpl 4.Sakacin_Q 5.BlpM
LP475	-
MAG.251	Lactacin_Fsubunit_lafA

Figures

(a)

(b)

Figure 1

(a) ANI analysis of different isolates of *L. pontis* (b) TNI analysis of different isolates of *L. pontis*

Figure 3

Phylogenetic tree constructed based on 1,281 core genes of five *L. pontis* strains and other three strains from same gene. The colours signify the origin of the strains: blue = fermented milk red; yellow = sourdough; orange = broiler chicken gastrointestinal tract

(a)

(b)

Figure 4

Annotation of the functional genomes of *L. pontis* based on the RAST database. (a): all RAST results (b): partial RAST results. A: cofactors, vitamins, prosthetic groups, pigments. B: cell wall and capsule. C: virulence, disease and defense. D: potassium metabolism. E: miscellaneous. F: phages, prophages, transposable elements, plasmids. G: membrane transport. H: iron acquisition and metabolism. I: RNA metabolism. J: nucleosides and nucleotides. K: protein metabolism. L: cell division and cell cycle. M:

regulation and cell signaling. N: DNA metabolism. O: fatty acids, lipids and isoprenoids. P: dormancy and sporulation. Q: respiration. R: stress responses. S: metabolism of aromatic compounds. T: amino acids and derivatives. U: sulfur metabolism. V: phosphorus metabolism. W: carbohydrates

Figure 5

Heatmap of Cazyase families of five strains of *L. pontis*. Blue = fermented milk strains. Red = sourdough strains. Yellow = broiler chicken gastrointestinal tract strain

Figure 6

Bacteriocins gene clusters present in *L. pontis* strains IMAU10341 and IMAU10345. (A) IMAU10341 has one core peptide. (B) IMAU10345 has one core peptide