MDPI

Article

Comparative Genomics of *Limosilactobacillus pontis* Strains: Niche-Specific Variations and Adaptations

Xueyan Lei [1,2,3,4], Qing Liu [1,2,3,4], Weicheng Li [1,2,3,4], Yu Li [1,2,3,4], Lixia Zhao [1,2,3,4] and Wenjun Liu [1,2,3,4,*]

[1]   Key Laboratory of Dairy Biotechnology and Engineering (IMAU), Ministry of Education, Inner Mongolia Agricultural University, Hohhot 010018, China; leixueyan2022@163.com (X.L.)
[2]   Key Laboratory of Dairy Products Processing, Ministry of Agriculture and Rural Affairs, Inner Mongolia Agricultural University, Hohhot 010018, China
[3]   Inner Mongolia Key Laboratory of Dairy Biotechnology and Engineering, Inner Mongolia Agricultural University, Hohhot 010018, China
[4]   Collaborative Innovative Center of Ministry of Education for Lactic Acid Bacteria and Fermented Dairy Products, Inner Mongolia Agricultural University, Hohhot 010018, China
*   Correspondence: wjliu168@163.com

Abstract: *Limosilactobacillus pontis* (*L. pontis*) is a species of lactic acid bacteria (LAB) found in various products, including fermented milk, sourdough, and broiler chickens' gastrointestinal tracts. However, the evolutionary strategies and genomic features of the species remain unknown, which limits its application. In this study, whole-genome sequencing was carried out to investigate the genomic characteristics and evolutionary strategies of *L. pontis*, combined with a comparative genomic approach, providing the necessary theoretical basis for its application. The genomic sequences in this study included three published genomic sequences and two strains isolated from fermented milk in our lab in Inner Mongolia, China. The mean genome size of the five *L. pontis* strains was 1.70 Mb, and the mean DNA G + C content was 53.06%. *L. pontis* was a species with higher G + C content in LAB. The phylogenetic evolutionary tree for the core genes showed that the strains from the same sources were aggregated into a cluster. There were some differences in the genomic information and phylogenetic relationships amongst *L. pontis* from different sources. An analysis of the annotation results identified differences in the functional genes, carbohydrate-active enzymes, and bacteriocins amongst different isolated strains, which were related to the environment. *L. pontis* was rich in glycosyltransferases, especially in strains isolated from fermented milk, indicating that they had higher sugar synthesis abilities. It is worth mentioning that only the *L. pontis* strains isolated in our laboratory identified the bacteriocin operon, which may increase its application potential. The genomic characteristics and evolutionary process of *L. pontis* were analyzed by comparative genomics, and this study explored the differences in the functional genes amongst the strains, aiming to provide new insights into the genetic characteristics and niche adaptations of *L. pontis*.

Keywords: *Limosilactobacillus pontis*; comparative genomics; niche adaptability

1. Introduction

Limosilactobacillus pontis (*L. pontis*) used to belong to *Lactobacillus* and is commonly found in fermenting animal and plant materials, such as sourdough, fermented milk, broiler chickens' gastrointestinal tracts, and vinegar [1]. It should be mentioned that because of the extreme phenotypic, ecological, and genotypic diversity among the species within the genus *Lactobacillus*, it has been reclassified into 25 genera, including the amended genera *Lactobacillus* and *Limosilactobacillus* and 23 others [2]. *Limosilactobacillus* includes *L. pontis*, *Limosilactobacillus reuteri*, and *Limosilactobacillus fermentum*, among others. Being generally recognized as safe [3], strains of this species have been widely applied in various foods' fermentation [4]. Kiran Shazadi [5] found that *L. reuteri* (MT180537) had a significant bactericidal effect on Enterococcus faecalis (MW051601) in the pharmaceutical industry; the

antimicrobial compounds or bacteriocin-like substances (BLIS) produced by these strains could be used in the preparation of ointments for intravaginal application against aerobic vaginitis (AV). Regarding *L. reuteri*, a member of the lactic acid bacteria (LAB) [6], a broad-spectrum antimicrobial substance is produced by this species during the fermentation of glycerol, called reuterin [7], and the secretion conditions include an appropriate temperature, weakly acidic conditions, and an anaerobic environment. It has been verified by experiments that reuterin is a mixture of monomers, hydrated monomers, and cyclic dimers of L-hydroxypropanal [8]. Reuterin has a large antibacterial range and can inhibit the growth of some Gram-positive and Gram-negative bacteria, fungi, and molds [9]. As it is in the same genus as *L. reuteri*, *L. pontis* may have potential bacteriocin-producing characteristics as well. Although *L. pontis* has been recorded from a wide range of sources, it has been the subject of few studies and is worthy of further exploration.

Comparative genomic studies of strains within the same species provide insights into genetic features, including gene modifications, gains, or losses, that facilitate the evolution and adaptation of strains to specific environmental niches [10]. In recent years, along with the development of rapid throughput DNA sequencing technologies, genome studies have become more feasible and affordable, and genomic data have become publicly available for numerous organisms [11]. However, the publicly available genomic data for *L. pontis* are very limited. Because pure cultures of *L. pontis* are not easy to obtain, there is little research on it. *L. pontis* is phylogenetically close to *Limosilactobacillus* vaginalis, *Limosilactobacillus oris*, and *L. reuteri* [1]. At present, compared with *L. pontis*, research on *L. reuteri* is favored in the field of *Limosilactobacillus*. *L. reuteri* can be isolated from the intestines of pigs, humans, and rodents. An investigation of *L. reuteri* from different sources has defined a phylogenetic tree based on core genomes that are distributed on different evolutionary branches, with habitats as clusters. This indicates that *L. reuteri* with different genetic backgrounds can selectively colonize and adapt to specific habitats, and this specific ability is significantly associated with the acid and bile salt tolerance and cell adhesion abilities of *L. reuteri* [12]. Most *L. reuteri* have good intestinal colonization and bacteriocin secretion abilities and have been widely used in daily life [13,14]. However, our understanding of *L. pontis'* genomic characteristics is poor, which limits its application; comparative genomics could be used to reveal its genomic characteristics.

We obtained two strains of *L. pontis* from fermented milk through the Lactic Acid Bacteria Collection Center (LABCC) of Inner Mongolia Agricultural University: *L. pontis* IMAU10341 and *L. pontis* IMAU10345. This study analyzed the genome sequences of *L. pontis* IMAU10341 and *L. pontis* IMAU10345 using Illumina Novaseq 6000 genome sequencing. We downloaded the type strain *L. pontis* DSM8475^T and three strains published by the NCBI (http://www.ncbi.nlm.nih.gov, accessed on 7 November 2023): *L. pontis* LP475, *L. pontis* UMB0683, and *L. pontis* MAG.251. We used a comparative genomics approach to define the pangenome, core genome, and unique genes of isolates from sourdough, fermented milk, and the broiler chicken gastrointestinal tract; assess the genetic diversity and gain insights into the distinguishing features; compare the numbers of functional genes and carbohydrate-active enzymes (CAZy); and explore the expression of bacteriocin genes. The aim of the current work was to carry out comparative genomics analyses of *L. pontis*, with more strains isolated from different niches, to explore its genetic diversity and the potential for host adaptation in this species.

2. Materials and Methods

2.1. Strains

The two strains evaluated in this study were *L. pontis* IMU10341 and *L. pontis* IMU10345 from the LABCC of Inner Mongolia Agricultural University; these strains had been isolated separately from two types of fermented milk collected in Dongwuzhumuqin Banner, Xilingol League, Inner Mongolia, China (N: 45°30′34.56″ E: 116°58′19.74″) [15]. The genomes of the remaining four *L. pontis* and 11 closely related strains were downloaded from the NCBI

Refseq (National Coalition Building Institute, https://www.ncbi.nlm.nih.gov/refseq/, accessed on 7 November 2023) database. The strain information is listed in Table 1.

Table 1. Genomic information for six strains of *L. pontis* and 11 closely related strains.

Strain	Isolation Source	CDS	Size (Mb)	G + C Content (%)	rRNA	tRNA	Accession Number
L. pontis IMAU10341	fermented milk	1787	1.77	52.49	6	64	GCF_026236575.1
L. pontis IMAU10345	fermented milk	1786	1.76	52.49	6	64	GCF_026236595.1
L. pontis DSM8475[T]	rye sourdough	1539	1.67	53.40	3	29	GCF_001435345.1
L. pontis LP475	rye sourdough	1561	1.71	53.30	18	67	GCF_009428965.1
L. pontis UMB0683	catheter	1698	1.87	48.70	7	59	GCF_002940945.1
L. pontis MAG.251	broiler chicken gastrointestinal tract	1551	1.60	53.60	1	32	GCF_017888165.1
L. antri DSM16041[T]	stomach mucosa	2038	2.30	51.00	2	45	GCF_000160835.1
L. oris DSM4864[T]	saliva	1861	2.03	50.00	2	51	GCF_001434465.1
L. panis DSM6035[T]	sourdough	1797	2.01	48.10	2	33	GCF_001435935.1
L. reuteri DSM20016[T]	intestine of adult	1889	2.00	38.90	18	68	GCF_000016825.1
L. vaginalis DSM5837[T]	vaginal swab	1681	1.79	40.50	3	24	GCF_001435915.1
L. equigenerosi DSM18793[T]	thoroughbred horses	1528	1.60	42.70	1	41	GCF_001435245.1
L. gastricus DSM16045[T]	gastric biopsies, stomach mucosa	1761	1.85	41.60	3	37	GCF_001434365.1
L. fermentum DSM20052[T]	fermented beets	1719	1.89	52.50	15	59	GCF_013394085.1
L. coleohominis DSM14060[T]	vagina	1715	1.72	40.80	3	29	GCF_001435055.1
L. ingluviei DSM15946[T]	crop	2055	2.15	49.90	2	69	GCF_001435775.1
L. mucosae DSM13345[T]	small intestine	1990	2.25	46.40	1	42	GCF_001436025.1

2.2. Main Instruments and Reagents

Genomic DNA extraction kits (Beijing Tiangen Biochemical Technology Co., Ltd., Beijing, China), MRS medium (Qingdao Haibo Biotechnology Co., Ltd., Qingdao, China), an electric heating constant-temperature incubator (Beijing Yiheng Technology Co., Ltd., Beijing, China), and an ultra-micro UV spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) [16].

2.3. Strain Culture and Genomic DNA Extraction

For each strain, an appropriate amount of freeze-dried bacterial powder was inoculated into 5 mL of MRS liquid medium and incubated anaerobically at 37 °C in the incubator for 24 h. Then, a 2% volume of inoculum was added to 15 mL of MRS liquid medium. The culture was incubated anaerobically for 24 h at 37 °C and the resulting bacteria collected by centrifugation [16]. A genomic DNA extraction kit was used to extract the genomic DNA, following the manufacturer's instructions, and the purity and quality of the DNA samples were checked with an ultra-micro ultraviolet spectrophotometer [17].

2.4. Genome Sequencing and Assembly

Sample quality was determined after the genomic DNA was extracted. Once the concentration and purity met the sequencing requirements, the Illumina Novaseq 6000 sequencing platform was used to complete whole-genome sequencing [18]. Moreover, 150 bp was selected to construct paired-end (PE) sequencing libraries and the average coverage of high-quality data was about 500× [19]. The quality of the obtained raw data was evaluated. After low-quality reads were filtered, adapters and primers were removed, the high-quality reads' sequences were obtained, and then the high-quality reads were spliced using the software SOAP denovo v2.04 (http://soap.genomics.org.cn/soapdenovo.html, accessed on 2 August 2021). Moreover, appropriate kmer values were selected [20]. Local inner gaps were then filled and single-base errors corrected using GapCloser v1.12 (http://soap.genomics.org.cn/soapdenovo.html, accessed on 3 August 2021) to finally complete the assembly of the genomes [21].

2.5. Comparative Genomics Analysis

2.5.1. Average Nucleotide Identity (ANI) Value Calculation

In this study, ANI values were calculated using FastANI (https://github.com/ParBLiSS/FastANI, accessed on 13 November 2023) according to Goris et al. [22,23]. TBtools (https://github.com/CJ-Chen/TBtools/releases, accessed on 13 November 2023) was used to draw an ANI heatmap [24].

2.5.2. Construction of Core Gene Set and Accessory Gene Set Analysis

After gene prediction based on the Prokka (v1.1.1) [25] software, the Roary (v3.13.0) [26] software was used to identify the core gene set and the accessory gene set, among which core genes were identified based on the principle that the amino acid similarity in the encoded protein was greater than 95%. A core gene is a gene that exists in the genomes of all strains; among them, a gene that appears in only one strain is a unique gene, and the unique genes from all of the strains represented the sum of the unique genes [27].

2.5.3. Phylogenetic Tree Construction

We used the core gene set obtained by Roary to construct a phylogenetic tree on the ITOL online website (https://itol.embl.de/, accessed on 14 November 2023) and explore the phylogenetic relationships among the *L. pontis* isolates [28].

2.5.4. Functional Gene Annotation

The genome sequences of all five *L. pontis* strains were annotated by uploading each sequence file to the Rapid Annotation using Subsystem Technology (RAST) web server (https://rast.nmpdr.org/rast.cgi, accessed on 17 November 2023) for genomic annotation. Protein coding sequences were predicted and categorized based on the SEED subsystem (most prominently FIGfams) in RAST [29].

2.5.5. CAZy Annotation

Carbohydrate-active enzymes were annotated for the genome sequences of the five *L. pontis* strains and compared and annotated through CAZy (http://www.cazy.org/, accessed on 22 November 2023). According to the results, the detailed carbohydrate-active enzyme family was searched on the CAZy webpage [30].

2.5.6. Bacteriocin

BAGEL4 (http://bagel4.molgenrug.nl/index.php, accessed on 27 November 2023) is an online database that facilitates the mining and visualization of ribosome-synthesized and post-translationally modified peptides and bacteriocin-producing gene clusters in the prokaryotic genome. On the BAGEL4 web server (accessed on 27 November 2023), a DNA nucleotide sequence was used as an input file [31].

2.6. Uploading Strain Genome

The assembled genomic data of the two strains (*L. pontis* IMAU10341 and *L. pontis* IMAU10345) were uploaded to the SRA database in the NCBI. The whole-genome shotgun sequences of strains IMAU10341 and IMAU10345 were deposited at DDBJ/EMBL/GenBank under the accession numbers JAIUWJ000000000 and JAIUWK000000000, respectively.

3. Results

3.1. General Genomic Characteristics

The G + C content of *L. pontis* UMB0683 was 48.70%, and the genome size was 1.87 Mb, which does not match the characteristics of the *L. pontis* genome and requires further analysis. Except for *L. pontis* UMB0683, the genome sizes of all strains of *L. pontis* ranged from 1.60 to 1.77 Mb, the average size was 1.78 Mb, and the average G + C content was 53.06%. The size of the assembled genome of *L. pontis* IMAU10341 was 1.77 Mb, encoding 1787 genes, and the G + C content was 52.49%; the size of the assembled genome of *L.*

pontis MAU10345 was 1.76 Mb, encoding 1786 genes, and the G + C 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 (Table 1).

3.2. ANI Analysis

The ANI is the gold standard for species identification, with ANI values of 95–96% being the boundary for species demarcation and an ANI value greater than 95% theoretically being considered to indicate the same species [22].

In this study, the ANI analysis was first performed on the whole-genome sequences of all strains in Table 1 and heat maps were drawn (Figure 1). The ANI values of these genomes were determined through pairwise comparison at the 95% threshold to further identify their species [22]. All strains clustered into two clusters: *L. oris* DSM4864[T], *Limosilactobacillus antri* DSM16041[T], *Limosilactobacillus panis* DSM6035[T], *L. reuteri* DSM20016[T], and *L. vaginalis* DSM5837[T] clustered in one large branch with *L. pontis*, indicating that they were more closely related to *L. pontis*; *L. pontis* IMAU10341, *L. pontis* IMAU10345, *L. pontis* LP475, and *L. pontis* MAG.251 clustered in one small branch with the type strain *L. pontis* DSM8475[T] (Figure 1). The ANI value of *L. pontis* DSM8475[T] was greater than 97%. In contrast, *L. pontis* UMB0683 was on the other small branch, and the ANI value with *L. pontis* DSM8475[T] was only 84% (less than 95%). In addition, the genomic characteristics of *L. pontis* UMB0683, downloaded from the NCBI, were also quite different from those of the other five strains. Thus, *L. pontis* UMB0683 is not a strain from this species and was excluded from further analyses. Then, the other five strains were analyzed, specifically *L. pontis* IMAU10341, *L. pontis* IMAU10345, *L. pontis* LP475, *L. pontis* DSM8475T, and *L. pontis* MAG.251.
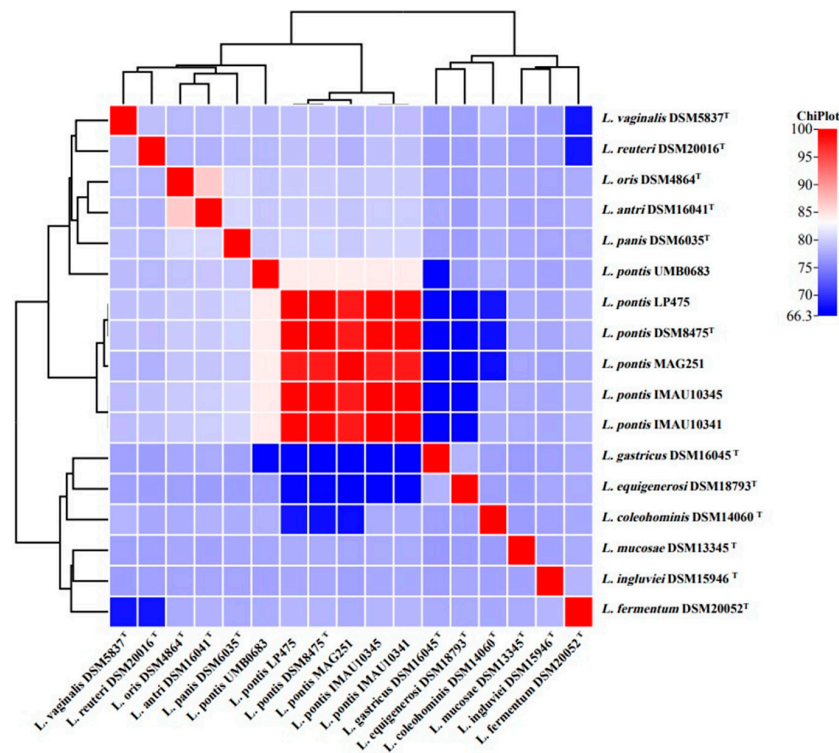

Figure 1. ANI analysis of *L. pontis* and closely related strains.

3.3. Core Genome and Pangenome Analysis

In order to compare the genetic differences between different strains, we used the Prokka (v1.1.1) and Roary (v3.13.0) software, with greater than 90% amino acid identity as the standard; the numbers of core genome and pangenome genes were counted [32]. The pangenome sets of the five strains of *L. pontis* contained 2166 genes, of which 1281 were core genome genes. There were eight and seven unique genes derived from fermented milk

strains IMAU10341 and IMAU10345, respectively, which were all pseudogenes; 37 and 52 unique genes in sourdough-derived strains DSM8475T and LP475, respectively; and 294 specific genes in the animal-derived strain MAG.251 (Figure 2).
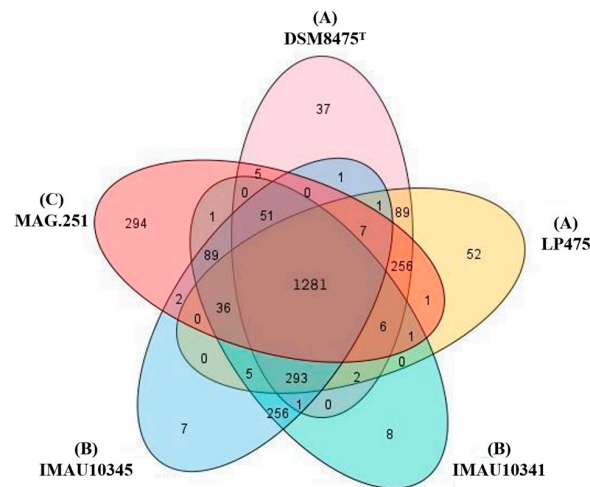

Figure 2. Venn diagram of core genomes of five *L. pontis* strains divided into three classes. (A) Sourdough, (B) fermented milk, (C) broiler chicken gastrointestinal tract.

3.4. Phylogenetic Tree

A phylogenetic tree directly shows the genetic distances and evolutionary relationships between individuals in the same population. Building a phylogenetic tree is a necessary prerequisite for the study of population structure and species evolution [33]. To further understand the population structure of *L. pontis*, we used neighbor joining to construct a phylogenetic tree, with a bootstrap value of 1000 [34]. This was based on comparisons between the nucleic acid sequences of the *L. pontis* strains and closely related species as outgroups: *L. oris* DSM4864T, *L. antri* DSM16041^T, *L. panis* DSM6035^T, *L. reuteri* DSM20016^T, and *L. vaginalis* DSM5837^T. The five strains of *L. pontis* were different from each other due to the development of genetic diversity during evolution; they were divided into three main branches (Figure 3). Isolates from sourdough (*L. pontis* LP475 and *L. pontis* DSM8475^T) were grouped together in one branch; strains from fermented milk (*L. pontis* IMAU10341 and *L. pontis* IMAU10345) were included in a second branch; and the branch containing the strain isolated from the broiler chicken gastrointestinal tract (*L. pontis* MAG.251) was closer to the root of the outgroup.
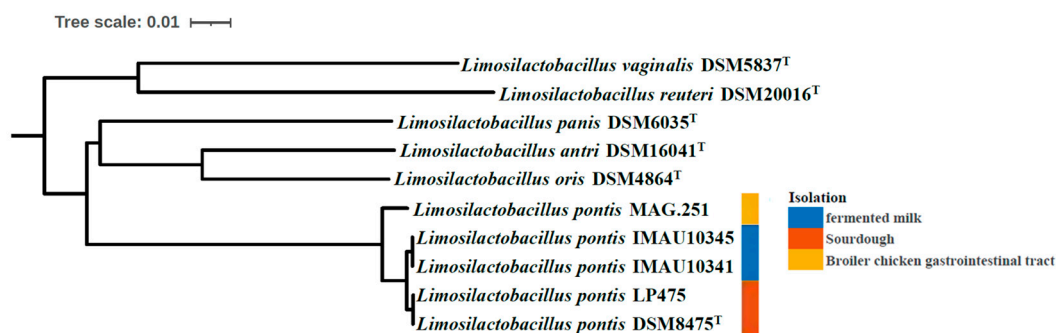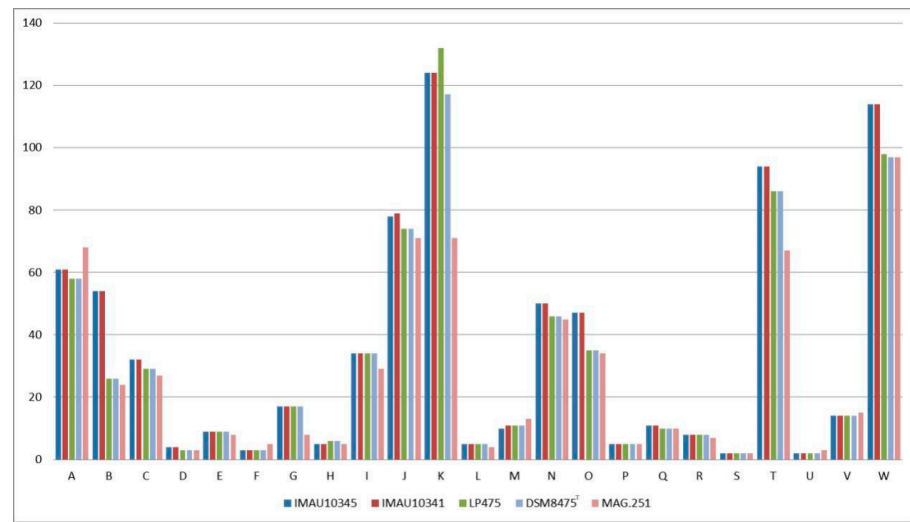

Figure 3. Phylogenetic tree constructed based on 279 core genes of five *L. pontis* strains and a further five closely related strains. The colors signify the origins of the strains. Blue = fermented milk, red = sourdough, yellow = broiler chicken gastrointestinal tract.
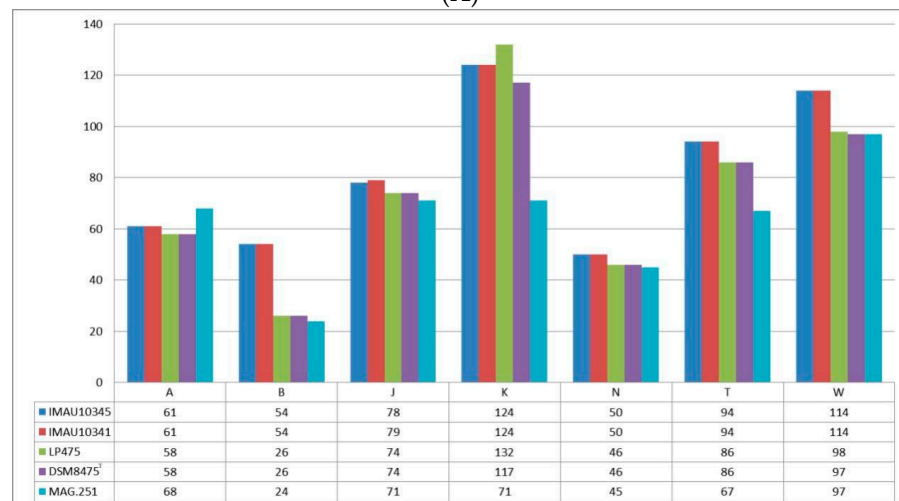
3.5. Functional Gene Analysis

The functional genes of the five *L. pontis* strains were predicted and annotated by RAST. A total of 23 functional categories were annotated (Figure 4). The *L. pontis* strains evaluated

had more genes related to the metabolism of nucleosides and nucleotides, proteins, amino acids and derivatives, and carbohydrates than any other type of gene. Strains *L. pontis* IMAU10341 and *L. pontis* IMAU10345, from fermented milk, had significantly more genes involved in the 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. The number of genes related to protein metabolism was highest in the sourdough isolate, *L. pontis* LP475. Genes related to cofactors, vitamins, prosthetic groups, and pigments were most numerous in the broiler chicken gastrointestinal tract strain, *L. pontis* MAG.251.

(**A**)

	A	B	J	K	N	T	W
IMAU10345	61	54	78	124	50	94	114
IMAU10341	61	54	79	124	50	94	114
LP475	58	26	74	132	46	86	98
DSM8475ᵀ	58	26	74	117	46	86	97
MAG.251	68	24	71	71	45	67	97

(**B**)

Figure 4. Annotation of *L. pontis*'s functional genes based on RAST database. (**A**) All RAST comment results and (**B**) some of the comment results. A: cofactors, vitamins, prosthetic groups, and pigments; B: cell wall and capsule; C: virulence, disease, and defense; D: potassium metabolism; E: miscellaneous; F: phages, prophages, transposable elements, and 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 response; S: metabolism of aromatic compounds; T: amino acids and derivatives; U: sulfur metabolism; V: phosphorus metabolism; W: carbohydrates.

3.6. Carbohydrate-Active Enzyme Analysis

To further study the differences in the genome of *L. pontis* and explore the reasons for the diversity in *L. pontis*, Prokka annotation and RAST annotation were performed for each strain. It was found that the proportion of carbohydrate-utilization-related functions in each genome was significantly different, and the number of related functional genes in the strains was quite different. Therefore, the abundance of the carbohydrate-related active enzymes amongst the strains was compared. CAZy is an enzyme database resource, based on the similarity of the amino acid sequences in the protein domains, which can synthesize or decompose various complex carbohydrates and sugar complexes [35]. The main function of the GH family is to hydrolyze or rearrange glycosidic bonds, including 152 gene families; the GT family is mainly responsible for the formation of glycosidic bonds, including 105 gene families; the function of the PL family is to break glycosidic bonds (not by hydrolysis), including 28 gene families; the CE family can hydrolyze carbohydrate esterases, including 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, including 83 gene families [36].

The CAZy annotation showing the distribution and abundance of CE, GH, and GT family genes in *L. pontis* was visualized as a heat map and clustered using the hierarchical clustering method. This showed that *L. pontis* was divided into three groups, representing the three different isolation sources. The predictions showed that the five strains contained 11 GH, eight GT, and one CE (Figure 5). There were differences in the composition of carbohydrate-active enzymes in the *L. pontis* strains isolated from different sources. *L. pontis* contained eight GT families, among which GT2 and GT4 were the most widely distributed, and GT2 was most abundant in strains from fermented milk. The GT32 family and GT101 family were only found in *L. pontis* MAG.251. *L. pontis* contained 11 GT families, of which GT36 and GT73 were the most widely distributed. GH53 was only found in sourdough strains. The functional genes varied in relation to the source and habitats, which may have resulted in evolution in different directions under the selection pressure of the living environment. GT2 and GT4 (glycosyltransferases, GTs), involved in the synthesis of cellulose and sucrose, were detected in L. pontis and were regarded as the largest family of carbohydrate enzymes, especially in strains isolated from fermented milk.
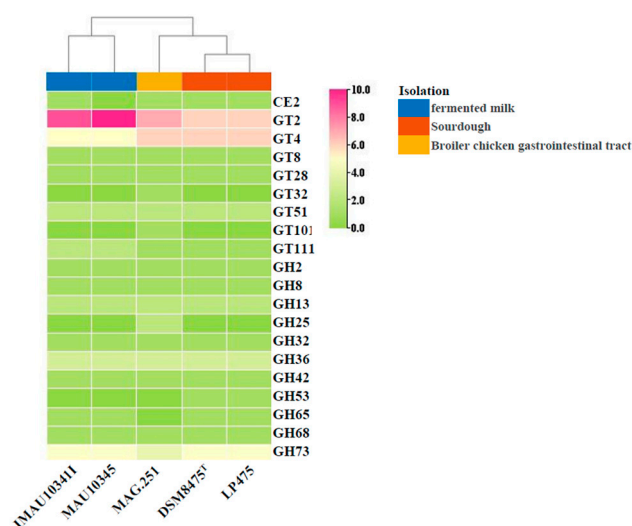

Figure 5. Heatmap of CAZyme families of *L. pontis* from different strains. Blue = fermented milk, red = sourdough, yellow = broiler chicken gastrointestinal tract.

3.7. Bacteriocin

In recent years, the use of software combined with databases has enabled us to quickly identify operons that can encode bacteriocins in the bacterial genome. Most studies use BAGEL to identify potential bacteriocin operons [37]. Bacteriocins can be roughly divided

into three classes; Class I bacteriocins encompass all peptides that undergo enzymatic modification during biosynthesis (heat-stable); Class II bacteriocins are unmodified bacteriocins (heat-stable); and Class III bacteriocins are large-molecular-weight and heat-labile antimicrobial proteins, usually composed of different domains [38].

We also used BAGEL to predict the bacteriocin operons in the five strains of *L. pontis*. The bacteriocin gene cluster prediction found that both *L. pontis* MAU10341 and *L. pontis* IMAU10345 from fermented milk contained the same core peptide, which was enterolysin A (Figure 6). Bacteriocin was not detected in the other three strains. Enterolysin A belongs to the Class III bacteriocins, is a novel antimicrobial protein, can degrade cell walls, and is mainly purified from an *Enterococcus faecalis* culture. Enterolysin A inhibits the growth of selected *Enterococci* [39]. The function of stationary-phase bacteriocin production might also be to kill bacteria in the same environment and thereby reduce the number of bacteria that compete for limiting nutrients [39]. Fermented milk is rich in nutrients and contains many types and a number of strains, so the strains need to produce bacteriocins for niche competition.
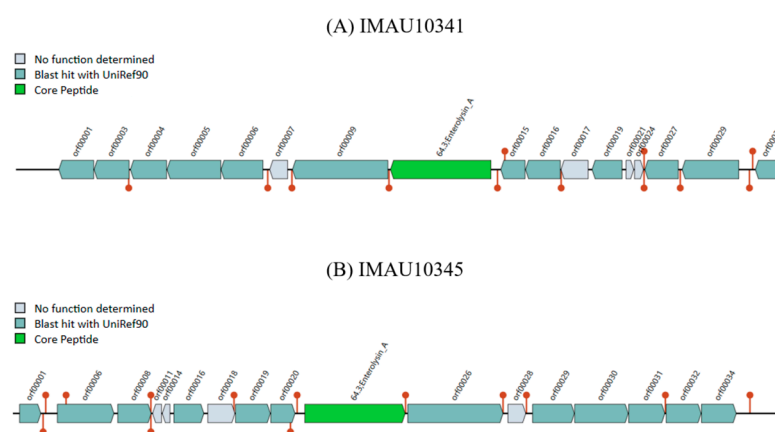

Figure 6. Bacteriocin gene clusters of IMAU10341 and IMAU10345. (**A**) IMAU10341 has one core peptide; (**B**) IMAU10345 has one core peptide.

4. Discussion

In this study, we undertook a comparative genomics analysis of the whole-genome sequences of five strains of *L. pontis* from different sources. It was found that *L. pontis* had a smaller genome and higher G + C content than those in the same genus. The length of the genome sequences of most *Lactobacillus* species is between 1.80 and 2.90 Mb, and the genome size of *L. pontis* was in the normal range. The G + C content of most *Lactobacillus* genomes is below 50%, and the G + C content of *Ligilactobacillus salivarius* UCC118 is at least 33% [39]. The G + C content of *L. pontis* in *Lactobacillus* was higher. The high G + C content of bacterial genomes results in a faster growth rate and a more stable DNA structure [40]. Not only is the genome density relatively large, but also the ability to resist high temperatures and alkaline environments is enhanced [41]. Compared with nomadic species and free-living species, the G + C content of the host-adapted Lactobacillus species was decreased, but the G + C content of Lactobacillus pontis was higher than that of other *Lactobacillus* species [42]. Based on the core gene phylogenetic tree, the fermented milk, sourdough, and animal isolates were placed in different branches, and the branch containing animal isolates was closer to the common ancestor. The conclusions from the ANI analysis were consistent with those of the phylogenetic tree analysis. The ANI values of *L. pontis* IMAU10341 and *L. pontis* IMAU10345 were high and different from the ANI value of the type strain. Pairs of strains isolated from similar samples were closely related, but not identical. Strains from fermented milk, sourdough, and the broiler chicken gastrointestinal tract naturally clustered into several categories. In addition, there were some differences in the ANI values between strains from different sources and the type strain, and the ANI values of *L. pontis*

strains from different hosts indicated that the species had sufficient genetic diversity to adapt to its different habitats.

In this study, the number of genes specific to a strain was similar when the strains were from the same source, but differed greatly when the strains were from different sources. The unique genes of *L. pontis* were under positive selection to adapt to different environments, and this study also demonstrated that food and intestinal habitats exerted different selective pressures related to the growth rate and metabolism [43]. The unique genes of *L. pontis* IMAU10341 and *L. pontis* IMAU10345 were pseudogenes. Some original genes of bacteria are degraded or inactivated if they are no longer needed. The formation of pseudogenes is one way in which genes decline or approach deletion from the genome [44]. Therefore, it is speculated that fermented milk, which is rich in readily available nutrition, represents a highly suitable environment for *L. pontis* strains and leads to the phenomenon of genome degradation in strains from fermented milk compared with strains from other sources.

Shuo Wang et al. [34] analyzed the phylogenetic relationships of *L. ruminis*, and a phylogenetic tree was created based on 91 orthologue genes that constituted the core genome. They found that strains from porcine, bovine, and human feces were divided into three clades; from the phylogenetic tree, it was clear that strains from the same sample type were clustered, closely related, and had the potential for niche adaptation. In our study, *L. pontis* MAG.251 was closest to the root, indicating that its genome was closer to its ancestors than the other strains evaluated. In different environments, the evolutionary process for species adaptation is different. The adaptation of bacterial genomes to specific environments will accumulate genetic variations related to the environmental characteristics. The genomes of the other four strains all changed due to environmental changes. Strains from different sources clustered together, indicating that the phylogenetic relationship reflected potential niche adaptations, which was consistent with the ANI results. The adaptive evolution of microorganisms to specific niches is also reflected in their functional genes, and specific niches will shape specific functional genes. In order to reduce the energy consumption of reproduction, strains living in specific niches will lose redundant functional genes related to the environment and simultaneously evolve specific genes adapted to these niches [45].

The functional gene annotation results showed that the *L. pontis* strains from fermented milk (*L. pontis* IMAU10341 and *L. pontis* IMAU10345) were rich in carbohydrate-related genes and genes involved in amino acids and derivatives. Genes related to the metabolism of fatty acids, lipids, and isoprenoids were significantly more abundant in fermented milk strains than in other strains, while genes for protein metabolism were more abundant in the sourdough strain *L. pontis* LP475 than in other strains. Based on RAST annotations, carbohydrates are the main energy source for the growth and development of organisms. Studies have shown that LAB can specifically adapt to the environment by obtaining carbohydrates related to the environment or losing non-essential carbohydrate genes. The differences among the five *L. pontis* strains in terms of their carbohydrate genes may be related to their different isolation sources; strains from the same isolation source were more similar to each other. Strains from fermented milk had more cell wall and capsule genes than other strains. This may be due to the rich environment and the greater number of microorganisms, which increase the thickness of the cell wall in order to resist external antibacterial substances. LAB require exogenous amino acids and peptides, which are provided by the hydrolysis of casein in fermented milk. Many LAB are deficient in a variety of amino acids [46]. In addition to enabling microbial growth, polypeptides, amino acids, and their derivatives also contribute to the distinctive texture and flavor of fermented milk, so the strains from fermented milk contained more genes associated with the metabolism of amino acids and their derivatives. These results also show that the functional genes of *L. pontis* in particular living environments function differently in their hosts.

In addition, all *L. pontis* strains contained genes for the production of extracellular polysaccharides, which can protect against external stresses. Because *L. pontis* contained more genes associated with carbohydrate metabolism, we also undertook a carbohydrate

enzyme analysis. There were differences in the composition of carbohydrate-active enzymes. *L. pontis* contained eight GT families, among which GT2 and GT4 were the most widely distributed. Studies have shown that GT2 and GT4 are rich in glycosyltransferases in naturally fermented milk. In fermented milk, these substances, especially polysaccharide compounds, are 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 had stronger sugar synthesis abilities. In addition, the GT32 family, which participates in the conversion of mannose, glucose, galactose, and others, and the GT101 family, which represents β-glucosyltransferase, were found only in *L. pontis* MAG.251, which may be related to its animal origin. GH53 was only found in sourdough isolates, which is related to its ability to degrade hemicellulose [13]. Carbohydrates are the main constituents of the cells and the main energy source and play a regulatory role in life activities [47]. Thus, the strains of *L. pontis* from different sources had different carbohydrate-active enzymes in order to adapt to their different habitats.

Bacteriocins are bacteriostatic substances produced by LAB in order to compete for ecological niches and enabling them to compete under real-life conditions [33]. Among the five strains of *L. pontis*, only those from fermented milk had predicted bacteriocins, which may be due to the rich nutritional content of fermented milk, leading to the diversity and abundance of other bacteria; bacteriocin is required for survival under niche competition.

The results of this study provide a reference for an understanding of the evolution of *L. pontis*. The functional genes of strains from different sources varied in order to adapt to various environments. *L. pontis* from fermented milk contained numerous carbohydrate-related genes and more genes for cell wall, capsule, and amino acid metabolism and its derivatives than strains from other sources. Protein metabolism and fatty acid, lipid, and isoprene metabolism genes were significantly more abundant in sourdough strains than other strains. In addition, all *L. pontis* strains contained genes for the production of extracellular polysaccharides, which protect from outside attacks. The two *L. pontis* strains from fermented milk exhibited the bacteriocin operon, which increases their application potential.

In order to adapt to different habitats, strains from different sources possess different phylogenetic relationships, functional genomic characteristics, carbohydrate enzymes, and bacteriocins. At present, there are few published genomes for *L. pontis*, which limits the in-depth study of the species, and it may be necessary to further explore the appropriate culture conditions and increase the number of strains from various sources and samples for research.

5. Conclusions

In this study, the genome sequences of five *L. pontis* strains provided a basis for an analysis of this species. *L. pontis* was characterized by high G + C content and potential niche adaptations. Moreover, the strains of *L. pontis* from different sources had a range of genetic diversity, which was closely related to their living environments; this lays a foundation for an exploration of the genetic basis and molecular evolution rules of *L. pontis*' adaptation to different habitats. However, due to the limited number of available strain sequences currently, more strain sequences need to be obtained for further comparative genomics studies to verify this. As more *L. pontis* strains are sequenced in the future, more *L. pontis* genomes from different environments can be compared, which will provide a strong indication of the factors affecting their genetic diversity and habitat adaptation.

Author Contributions: Conceptualization, X.L. and Q.L.; methodology, X.L., Q.L. and W.L. (Weicheng Li); formal analysis, W.L. (Weicheng Li), Y.L. and L.Z.; writing—review and editing, X.L.; resources, W.L. (Wenjun Liu); data curation, W.L. (Wenjun Liu). All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Vogel, R.F.; BöCKER, G.; Stolz, P.; Ehrmann, M.; Fanta, D.; Ludwig, W.; Pot, B.; Kersters, K.; Schleifer, K.H.; Hammes, W.P. Identification of *Lactobacilli* from sourdough and description of *Lactobacillus pontis* sp. nov. *Int. J. Syst. Evol. Microbiol.* **1994**, *44*, 223–229. [CrossRef]

2. Zheng, J.; Wittouck, S.; Salvetti, E.; Franz, C.M.; Harris, H.M.; Mattarelli, P.; O'toole, P.W.; Pot, B.; Vandamme, P.; Walter, J. 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.* **2020**, *70*, 2782–2858. [CrossRef]

3. Maya-Barrios, A.; Lira-Hernandez, K.; Jiménez-Escobar, I.; Hernández, L.; Ortiz-Hernandez, A.; Jiménez-Gutiérrez, C.; López-Velázquez, G.; Gutiérrez-Castrellón, P. *Limosilactobacillus reuteri* ATCC PTA 5289 and DSM 17938 as adjuvants to improve evolution of pharyngitis/tonsillitis in children: Randomised controlled trial. *Benef. Microbes* **2021**, *12*, 137–145. [CrossRef]

4. Bao, Y.; Zhang, Y.; Zhang, Y.; Liu, Y.; Wang, S.; Dong, X.; Wang, Y.; Zhang, H. Screening of potential probiotic properties of *Lactobacillus fermentum* isolated from traditional dairy products. *Food Control* **2010**, *21*, 695–701. [CrossRef]

5. Shazadi, K.; Arshad, N. Evaluation of inhibitory and probiotic properties of lactic acid bacteria isolated from vaginal microflora. *Folia Microbiol.* **2022**, *67*, 427–445. [CrossRef]

6. Saviano, A.; Brigida, M.; Migneco, A.; Gunawardena, G.; Zanza, C.; Candelli, M.; Franceschi, F.; Ojetti, V. *Lactobacillus Reuteri* DSM 17938 (*Limosilactobacillus reuteri*) in diarrhea and constipation: Two sides of the same coin? *Medicina* **2021**, *57*, 643. [CrossRef] [PubMed]

7. Talarico, T.; Casas, I.; Chung, T.C.; Dobrogosz, W. Production and isolation of reuterin, a growth inhibitor produced by *Lactobacillus* reuteri. *Antimicrob. Agents Chemother.* **1988**, *32*, 1854–1858. [CrossRef]

8. Arqués, J.L.; Fernández, J.; Gaya, P.; Nuñez, M.; Rodríguez, E.; Medina, M. Antimicrobial activity of reuterin in combination with nisin against food-borne pathogens. *Int. J. Food Microbiol.* **2004**, *95*, 225–229. [CrossRef] [PubMed]

9. Dobrogosz, W.J.S.E. 5413960 Antibiotic reuterin: Dobrogosz Walter J.; Lindgren Sven E Raleigh, NC, United States Assigned to Biogaia AB. *Biotechnol. Adv.* **1996**, *14*, 288. [CrossRef]

10. Feyereisen, M.; Mahony, J.; Kelleher, P.; Roberts, R.J.; O'Sullivan, T.; Geertman, J.-M.A.; van Sinderen, D. Comparative genome analysis of the *Lactobacillus brevis* species. *BMC Genom.* **2019**, *20*, 416. [CrossRef]

11. Francis, F.; Kim, J.; Ramaraj, T.; Farmer, A.; Rush, M.C.; Ham, J.H. Comparative genomic analysis of two Burkholderia glumae strains from different geographic origins reveals a high degree of plasticity in genome structure associated with genomic islands. *Mol. Genet. Genom.* **2013**, *288*, 195–203. [CrossRef]

12. Oh, P.L.; Benson, A.K.; Peterson, D.A.; Patil, P.B.; Moriyama, E.N.; Roos, S.; Walter, J. Diversification of the gut symbiont *Lactobacillus reuteri* as a result of host-driven evolution. *ISME J.* **2010**, *4*, 377–387. [CrossRef]

13. Jiang, J.Y.; Li, K.; Xiao, Y.; Zhong, A.; Tang, J.J.; Duan, Y.F.; Li, Z.J. *Limosilactobacillus reuteri* Regulating Intestinal Function: A Review. *Fermentation* **2022**, *9*, 19. [CrossRef]

14. Tyagi, A.; Yeon, S.J.; Chelliah, R.; Oh, D.H. Draft genome sequence of *Limosilactobacillus reuteri*, isolated from human breast milk. *Microbiol. Resour. Announc.* **2023**, *12*, 112822. [CrossRef]

15. Liu, W.; Bao, Q.; Jirimutu; Qing, M.; Siriguleng; Chen, X.; Sun, T.; Li, M.; Zhang, J.; Yu, J. Isolation and identification of lactic acid bacteria from Tarag in Eastern Inner Mongolia of China by 16S rRNA sequences and DGGE analysis. *Microbiol. Res.* **2012**, *167*, 110–115. [CrossRef]

16. Yu, J.; Zhao, J.; Song, Y.; Zhang, J.; Yu, Z.; Zhang, H.; Sun, Z. Comparative genomics of the herbivore gut symbiont *Lactobacillus reuteri* reveals genetic diversity and lifestyle adaptation. *Front. Microbiol.* **2018**, *9*, 1151. [CrossRef]

17. Tanaka, T.; Sakai, R.; Kobayashi, R.; Hatakeyama, K.; Matsunaga, T. Contributions of phosphate to DNA adsorption/desorption behaviors on aminosilane-modified magnetic nanoparticles. *Langmuir* **2009**, *25*, 2956–2961. [CrossRef]

18. Modi, A.; Vai, S.; Caramelli, D.; Lari, M. The Illumina Sequencing Protocol and the NovaSeq 6000 System. *Methods Mol. Biol.* **2021**, *2242*, 15–42. [CrossRef]

19. Caporaso, J.G.; Lauber, C.L.; Walters, W.A.; Berg-Lyons, D.; Huntley, J.; Fierer, N.; Owens, S.M.; Betley, J.; Fraser, L.; Bauer, M.; et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **2012**, *6*, 1621–1624. [CrossRef]

20. Luo, R.; Liu, B.; Xie, Y.; Li, Z.; Huang, W.; Yuan, J.; He, G.; Chen, Y.; Pan, Q.; Liu, Y. SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *GigaScience* **2012**, *1*, 2047-217X-1-18. [CrossRef]

21. Xu, M.; Guo, L.; Gu, S.; Wang, O.; Zhang, R.; Peters, B.A.; Fan, G.; Liu, X.; Xu, X.; Deng, L. TGS-GapCloser: A fast and accurate gap closer for large genomes with low coverage of error-prone long reads. *GigaScience* **2020**, *9*, 94. [CrossRef]

22. Goris, J.; Konstantinidis, K.T.; Klappenbach, J.A.; Coenye, T.; Vandamme, P.; Tiedje, J.M. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* **2007**, *57*, 81–91. [CrossRef]

23. Chen, J.; Yang, X.; Chen, J.; Cen, Z.; Guo, C.; Jin, T.; Cui, Y. SISP: A Fast Species Identification System for Prokaryotes Based on Total Nucleotide Identity of Whole Genome Sequences. *Infect. Dis. Transl. Med.* **2015**, *1*, 30–55. [CrossRef]

24. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202. [CrossRef]

25. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [CrossRef]

26. Page, A.J.; Cummins, C.A.; Hunt, M.; Wong, V.K.; Reuter, S.; Holden, M.T.; Fookes, M.; Falush, D.; Keane, J.A.; Parkhill, J. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **2015**, *31*, 3691–3693. [CrossRef]

27. Duranti, S.; Milani, C.; Lugli, G.A.; Mancabelli, L.; Turroni, F.; Ferrario, C.; Mangifesta, M.; Viappiani, A.; Sánchez, B.; Margolles, A. Evaluation of genetic diversity among strains of the human gut commensal *Bifidobacterium adolescentis*. *Sci. Rep.* **2016**, *6*, 23971. [CrossRef]

28. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [CrossRef]

29. Aziz, R.K.; Bartels, D.; Best, A.A.; DeJongh, M.; Disz, T.; Edwards, R.A.; Formsma, K.; Gerdes, S.; Glass, E.M.; Kubal, M. The RAST Server: Rapid annotations using subsystems technology. *BMC Genom.* **2008**, *9*, 75. [CrossRef]

30. Cantarel, B.L.; Coutinho, P.M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res.* **2009**, *37*, 233–238. [CrossRef]

31. van Heel, A.J.; de Jong, A.; Song, C.; Viel, J.H.; Kok, J.; Kuipers, O.P. BAGEL4: A user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res.* **2018**, *46*, 278–281. [CrossRef]

32. Johnson, B.; Shneiderman, B. *Tree-Maps: A Space-Filling Approach to the Visualization of Hierarchical Information Structures*; IEEE: San Diego, CA, USA, 1991; pp. 284–291. [CrossRef]

33. Wang, S.; Yang, B.; Ross, R.P.; Stanton, C.; Zhao, J.; Zhang, H.; Chen, W. Comparative genomics analysis of *Lactobacillus ruminis* from different niches. *Genes* **2020**, *11*, 70. [CrossRef]

34. Balakirev, E.S.; Krupnova, T.N.; Ayala, F.J. DNA variation in the phenotypically-diverse brown alga Saccharina japonica. *BMC Plant Biol.* **2012**, *108*, 1471–2229. [CrossRef]

35. Accetto, T.; Avguštin, G. Polysaccharide utilization locus and CAZYme genome repertoires reveal diverse ecological adaptation of Prevotella species. *Syst. Appl. Microbiol.* **2015**, *38*, 453–461. [CrossRef]

36. Cao, H.; Ekstrom, A.; Yin, Y. *Plant Carbohydrate Active Enzyme (CAZyme) Repertoires: A Comparative Study*; Springer International Publishing: New York, NY, USA, 2015; pp. 115–134. [CrossRef]

37. Vilma, K.; Mantas, S.; Kristina, B.; Naujokaitytė, G.; Mulkytė, K.; Malakauskas, M.; Maruška, A. Current state of purification, isolation and analysis of bacteriocins produced by lactic acid bacteria. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 1323–1335. [CrossRef]

38. Alvarez-Sieiro, P.; Montalban-Lopez, M.; Mu, D.; Kuipers, O. Bacteriocins of lactic acid bacteria: Extending the family. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 2939–2951. [CrossRef]

39. Sinéad, C.; Li, Y.; Riedel, C.U.; O'Toole, P.W.; Hill, C.; Gahan, C.G.M. From the Cover: Bacteriocin production as a mechanism for the antiinfective activity of *Lactobacillus salivarius* UCC118. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 761. [CrossRef]

40. Bobay, L.-M.; Ochman, H. Impact of recombination on the base composition of bacteria and archaea. *Mol. Biol. Evol.* **2017**, *34*, 2627–2636. [CrossRef]

41. Musto, H.; Naya, H.; Zavala, A.; Romero, H.; Alvarez-Valín, F.; Bernardi, G. Correlations between genomic GC levels and optimal growth temperatures in prokaryotes. *FEBS Lett.* **2004**, *573*, 73–77. [CrossRef]

42. Xiao, Y.; Zhao, J.; Zhang, H.; Zhai, Q.; Chen, W. Mining genome traits that determine the different gut colonization potential of *Lactobacillus* and *Bifidobacterium* species. *Microb. Genom.* **2021**, *7*, 000581. [CrossRef]

43. Maldonado-Gómez, M.X.; Martínez, I.; Bottacini, F.; O'Callaghan, A.; Ventura, M.; van Sinderen, D.; Hillmann, B.; Vangay, P.; Knights, D.; Hutkins, R.W. Stable engraftment of *Bifidobacterium longum* AH1206 in the human gut depends on individualized features of the resident microbiome. *Cell Host Microbe* **2016**, *20*, 515–526. [CrossRef]

44. Lerat, E.; Ochman, H. Recognizing the pseudogenes in bacterial genomes. *Nucleic Acids Res.* **2005**, *33*, 3125–3132. [CrossRef]

45. Jung, J.; Kim, K.; Yoo, D.; Lee, C.; Kang, J.; Cho, K.; Kang, D.-K.; Kwak, W.; Yoon, S.H.; Sohn, H. Comparative genomic analysis of *Lactobacillus plantarum* GB-LP4 and identification of evolutionarily divergent genes in high-osmolarity environment. *Genes Genom.* **2018**, *40*, 217–223. [CrossRef]

46. Walter, J.; Britton, R.A.; Roos, S. Host-microbial symbiosis in the vertebrate gastrointestinal tract and the *Lactobacillus reuteri* paradigm. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 4645–4652. [CrossRef]

47. El Kafsi, H.; Binesse, J.; Loux, V.; Buratti, J.; Boudebbouze, S.; Dervyn, R.; Kennedy, S.; Galleron, N.; Quinquis, B.; Batto, J.M. Lactobacillus delbrueckii ssp. *lactis* and ssp. *bulgaricus*: A chronicle of evolution in action. *BMC Genom.* **2014**, *15*, 407. [CrossRef]