

Article

Comparative Genomic Analysis of *Lactiplantibacillus plantarum*: Insights into Its Genetic Diversity, Metabolic Function, and Antibiotic Resistance

Ruiqi Li ¹ and Chongpeng Bi ^{2,*}

¹ College of Veterinary Medicine, Northeast Agricultural University, Harbin 150030, China; 13034591915@163.com

² College of Animal Science and Technology, Northeast Agricultural University, Harbin 150030, China

* Correspondence: bnm0722@163.com

Abstract

Background/Objectives: *Lactiplantibacillus plantarum* is widely utilized in the fermentation industry and offers potential health benefits. However, large-scale comparative genomic analyses aimed at exploring its metabolic functions and conducting safety assessments are still lacking. **Methods:** In this study, we performed a comparative genomic analysis of 324 *L. plantarum* strains sourced from various origins and geographical locations. **Results:** The results revealed that *L. plantarum* possesses a total of 2403 core genes, of which 12.3% have an unknown function. The phylogenetic analysis revealed a mixed distribution from various origins, suggesting complex transmission pathways. The metabolic analysis demonstrated that *L. plantarum* strains can produce several beneficial metabolites, including lysine, acetate, and riboflavin. Furthermore, *L. plantarum* is highly capable of degrading various carbohydrates and proteins, increasing its adaptability. Further, we profiled the antimicrobial peptides (AMPs) in the genomes of *L. plantarum*. We identified a widely distributed AMP and its variants, presenting in a total of 280 genomes. In our biosafety assessment of *L. plantarum*, we identified several antibiotic resistance genes, such as *Tet(M)*, *ANT(6)-Ia*, and *mdeA*, which may have potential for horizontal gene transfer within the *Lactobacillaceae* family. **Conclusions:** This study provides genomic insights into the genetic diversity, metabolic functions, antimicrobial properties, and biosafety of *L. plantarum*, underscoring its potential applications in biotechnology and environmental adaptation.

Keywords: *Lactiplantibacillus plantarum*; genomic analysis; metabolic function; probiotic mechanism; biosafety



Academic Editor: Silvia Turrone

Received: 30 June 2025

Revised: 19 July 2025

Accepted: 22 July 2025

Published: 24 July 2025

Citation: Li, R.; Bi, C. Comparative Genomic Analysis of *Lactiplantibacillus plantarum*: Insights into Its Genetic Diversity, Metabolic Function, and Antibiotic Resistance. *Genes* **2025**, *16*, 869. <https://doi.org/10.3390/genes16080869>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Lactiplantibacillus plantarum, formerly known as *Lactobacillus plantarum*, is a species of lactic acid bacteria that can be isolated from a variety of sources, including humans, animals, grass silage, kimchi, pickled cabbage, cheese, etc. [1]. This bacterium can metabolize pentose and hexose sugars to produce lactic acid, carbon dioxide, and either acetate or ethanol [2]. Thus, it plays a crucial role in the fermentation of foods such as sauerkraut, kimchi, pickles, and dairy products [3]. In addition to being an important component in food fermentation, *L. plantarum* also functions as a probiotic, enhancing human health. For example, *L. plantarum* L168 has been shown to have several beneficial properties, including restoring the structure of the gut microbiota, modulating the concentration of the neurotransmitter serotonin, and improving impaired behavior [4].

The widespread application of *L. plantarum* in the food industry and its potential as a probiotic necessitate a comprehensive understanding of its genetic diversity, phylogenetic relationships, and functional characteristics. In recent years, advancements in next-generation sequencing technologies have facilitated numerous genomic studies of *L. plantarum* strains, providing deep insights. In the genome of *L. plantarum* LP-F1, a total of 130 enzymes related to carbohydrate metabolism were identified, highlighting its significant capacity for carbohydrate utilization [5]. A genomic analysis of *L. plantarum* L125 revealed the presence of an incomplete plataricin gene cluster, with no bacteriocin-like activity detected [6]. Additionally, a genomic analysis of *L. plantarum* BRD3A suggested its potential to produce bacteriocins such as plantaricin E, lantarin F, and enterocin X [7]. Furthermore, the genomic analyses also indicated a higher prevalence of antimicrobial resistance genes in 212 *L. plantarum* genomes, particularly those associated with tetracycline-resistance [8].

With the accumulation of *L. plantarum* genomes, a comparative genomic analysis would be helpful in providing a comprehensive overview of their characteristics. Garcia-Gonzalez and colleagues conducted a comparative genomic analysis of 42 *L. plantarum* strains to gain insights into their probiotic properties [9]. Similarly, Evanovich and colleagues also performed a study that did not identify any antibiotic resistance genes (ARGs) in the *L. plantarum* genome [10]. A recent investigation evaluated the antibiotic resistance patterns of *L. plantarum* strains, revealing the presence of resistance to ampicillin [11]. Martino and colleagues noted a loose connection between the distribution and source of *L. plantarum* [12], which may be attributed to the insufficient number of genomes available. With the increasing number of genomes, there is an opportunity for large-scale comparative genomic analyses to reevaluate these relationships.

In this study, we collected a total of 324 complete genomes of *L. plantarum* strains from NCBI, which were isolated from various sources across six continents. A comprehensive genomic analysis was conducted to explore the genomic characteristics, phylogenetic relationships, metabolic functions, and biosafety of these strains. The findings of this study provide valuable insights into the probiotic properties and biosafety of *L. plantarum*, offering a scientific basis for its application in the food industry and probiotic development.

2. Materials and Methods

2.1. Collection of Microbial Genomes

To better understand the genomic characteristics of *L. plantarum*, a total of 324 complete genomes (as of 29 March 2025) were deposited in the National Center for Biotechnology Information (NCBI, Supplementary Table S1). To further confirm the quality of these genomes, we utilized GTDB-Tk v2.4.0 [13] and checkM2 v1.0.1 [14] to verify the taxonomic classification and genome quality, respectively. The phylogenetic relationships among different strains and plasmids were calculated using average nucleotide identity (ANI) with pyani v0.2.3 [15]. These genomes were subsequently annotated with Prokka v1.14.6 [16], which identified open reading frames for downstream analyses.

2.2. Construction of the Phylogenetic Tree

A pan-genome analysis of the 324 genomes was conducted using Panaroo v1.14.6 [17] to explore the genetic diversity and identify core and accessory genes within the *L. plantarum* strains. The functional annotations of the core genes were performed using eggNOG v2.1.12 [18].

We then constructed the phylogenetic tree using IQ-TREE2 v2.1.4-beta [19], which is based on the core genome alignment with the parameters of “-m TEST -B 1000 -bnni”. The phylogenetic tree was further clustered using fastBAPS [20].

2.3. Prediction of Metabolic Profiles

Functional annotations of the *L. plantarum* genomes were conducted using the metabolic-G.pl program in METABOLIC v4.0 [21], which integrates the Kyoto Encyclopedia of Genes and Genomes (KEGG), TIGRfam, Pfam, dbCAN2, and MEROPS databases. This process includes a protein motif validation step to evaluate the presence of metabolic pathways on the basis of KEGG modules. The annotations were subsequently manually reviewed for the presence of KEGG modules, carbohydrate-active enzymes (CAZymes), peptidases, and their inhibitors.

2.4. In Silico Analysis of Virulence Genes and Antimicrobial Resistance and Their Associated Mobile Genetic Elements (MGEs)

Virulence genes were annotated using ABRicate (<https://github.com/tseemann/abricate> (accessed on 6 March 2025)) against the Virulence Factor Database using the default parameters [22]. ARGs were annotated using the Comprehensive Antibiotic Resistance Database (CARD, v6.0.2) [23]. To reduce the number of potential false positives, only perfect and strict hits were retained. To establish links between these genes and MGEs, a detailed analysis of the *L. plantarum*-associated prophage fragments and plasmids was conducted. First, we used geNomad v1.5.2 [24] and VITAP v1.7.1 [25] to identify the prophage fragments and perform the taxonomic classification. We then typed these plasmids using the function “mob_typer” embedded in the software MOB-suite v3.1.0 [26].

2.5. In Silico Analysis of Antimicrobial Peptides (AMPs)

Macrel v1.3.0 [27] was run directly on the genomes to identify the AMPs. To assess the novelty of the predicted AMPs, we manually constructed an AMP database comprising 44,406 AMPs sourced from five existing AMP databases [28–32]. Subsequently, the identified AMPs were queried against the curated AMP database with blastp v2.10.0+ [33]. Nonredundant AMP sequences were then clustered using CD-HIT v4.8.1 [34], with an identity of over 70%. The amino acid sequences of AMPs, which belong to the same cluster, were aligned using MUSCLE v3.8.31 with the default parameters [35]. The alignments were then used to construct phylogenetic trees using IQ-TREE2. To better understand the potential effects of the changes to the AMP sequences, APEX [36] was used to predict the species-specific antimicrobial activities of the AMPs, which were measured by the minimum inhibitory concentration (MIC) against 34 type strains.

2.6. Statistics and Visualization

The data in this study were expressed as the mean $\pm$ SD. The phylogenetic trees were visualized using iTOL v7 [37]. Circular genome maps of ARG-carrying plasmids were visualized using Proksee [38]. The other visualizations were constructed using R software v4.3.1, primarily with the pheatmap v1.0.12 [39] and ggplot2 v3.4.4 [40] packages. Pairwise comparisons between different groups were performed using the “pairwise.wilcoxon.test” function in R.

3. Results

3.1. Strain Information and Genome Characteristics

In this study, we collected a total of 324 complete *L. plantarum* genome sequences from NCBI (Supplementary Table S1). These strains were isolated from a variety of sources, including human-related samples ($n = 62$), fermented foods ($n = 138$), animal-related samples ($n = 56$), plants ($n = 26$), the environment ($n = 7$), and other sources ($n = 35$). The samples were distributed across six continents: Asia, Europe, North America, South America, Oceania, and Africa. These genome sizes ranged from 2,793,376 bp to 3,687,306 bp,

with an average length of 3,305,355 bp. The genomes contained an average of 2.9 ± 2.8 plasmids, accounting for an average of $101,908 \pm 98,232$ bp.

The pan-genome analysis revealed that these genomes contained a total of 12,416 genes, of which 2403 were classified as core genes, including both core genes and soft-core genes (Table 1). The five most common COG categories of the core genes were classified as the following: [S] function unknown (12.3%); [K] transcription (6.0%); [E] amino acid transport and metabolism (5.5%); [J] translation, ribosomal structure, and biogenesis (4.6%); and [G] carbohydrate transport and metabolism (4.5%).

Table 1. Number of core and accessory genes in 324 genomes.

Category	Cut-Off	Number of Genes
Core genes	$99\% \leq \text{strains} \leq 100\%$	2194
Soft-core genes	$95\% \leq \text{strains} < 99\%$	209
Shell genes	$15\% \leq \text{strains} < 95\%$	1164
Cloud genes	$0\% < \text{strains} < 15\%$	8849
Total genes	$0\% < \text{strains} \leq 100\%$	12,416

3.2. Phylogenetic Analysis Reveals Complex Transmission Pathways

To investigate the phylogenetic relationships among the *L. plantarum* strains, a phylogenetic tree was constructed on the basis of the core genome, and the strains were grouped into seven different clusters (Figure 1). Three larger clusters were observed in the phylogenetic tree, consisting of 188 genomes in cluster 1, 82 genomes in cluster 7, and 39 genomes in cluster 2. From a host perspective, these three larger clusters included the strains from all five specific hosts included in this study, indicating no preference for any particular host (Supplementary Figure S1). From a geographical perspective, strains from other continents were distributed sporadically across the three larger clusters. This suggests a wider distribution of these genetically related strains. Additionally, there were no obvious associations between the host and the cluster, potentially because of the frequent transmission among various origins.

3.3. Key Metabolic Functions Revealed by the Presence of KEGG Modules in *L. plantarum*

Different strains of the same species may exhibit different metabolic functions and modulate host physiology by producing various metabolites. To better understand the potential production of metabolites, we further explored their metabolic functions by analyzing the presence of the KEGG modules. A total of 41 KEGG modules, belonging to 17 categories, were found in the 324 genomes (Supplementary Figure S2). These KEGG modules were involved mainly in the following categories: central carbohydrate metabolism ($n = 8$), cofactor and vitamin metabolism ($n = 7$), purine metabolism ($n = 4$), other carbohydrate metabolism ($n = 4$), arginine and proline metabolism ($n = 4$), etc. A total of 30 out of the 41 KEGG modules were found in most strains ($\geq 95\%$). The presence of these modules indicates that *L. plantarum* can produce various metabolites, including lysine (M00525), acetate (M00579), inosine monophosphate (IMP, M00048), riboflavin (M00125), and tetrahydrofolate (M00126). And a large proportion of *L. plantarum* strains ($n = 304$, 93.8%) could also produce histidine (M00026), a conditionally essential amino acid, suggesting the existence of heterogeneity.

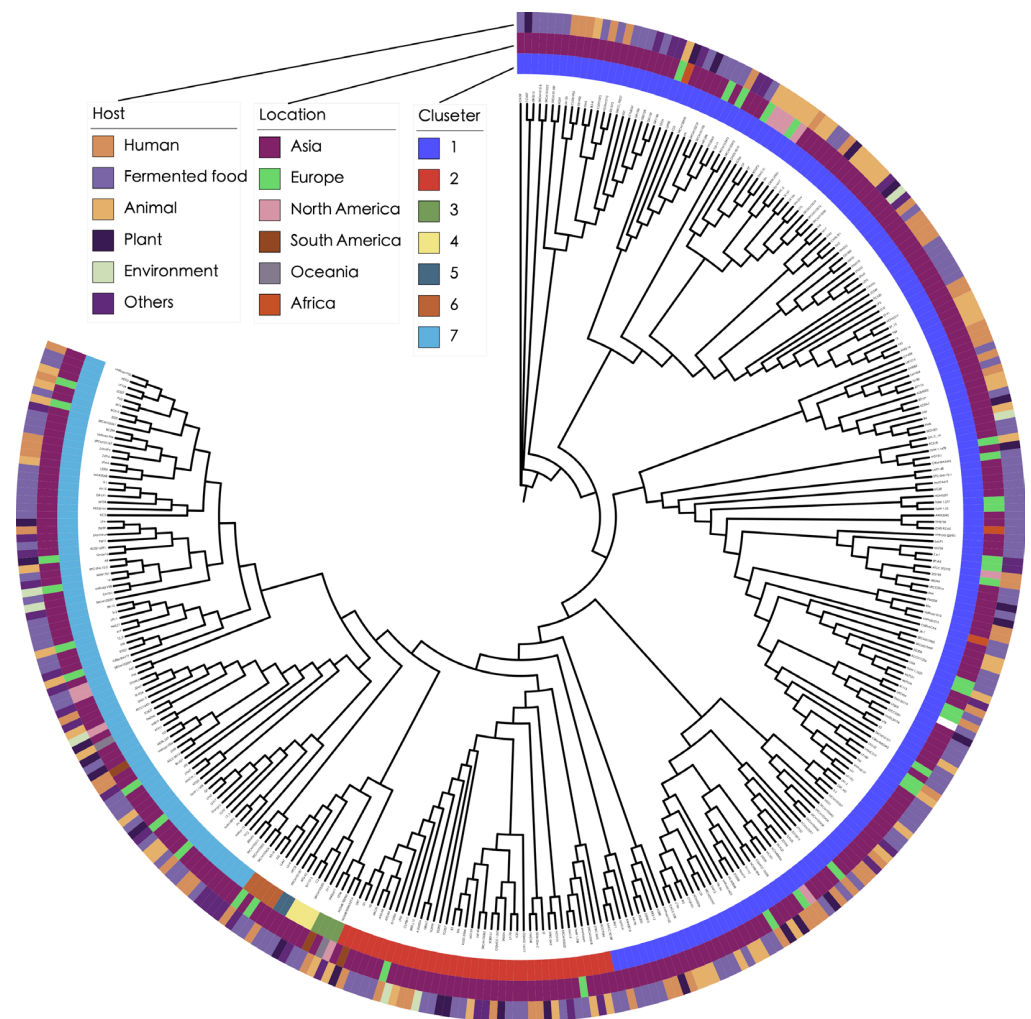


Figure 1. Phylogenetic tree constructed based on the core genomes of 324 *L. plantarum* strains. From the inner to outer, the three rings represent different clusters, the location, and the host of strains.

We then compared the differences in the KEGG modules between strains from various origins. This revealed that the KEGG modules of human-derived *L. plantarum* were similar to those of *L. plantarum* from animal and fermented food sources. Compared with the environmental samples (100.0%), human-, animal-, and fermented food-derived samples presented lower proportions of C1-unit interconversion (M00140), ranging from 66.1% to 76.8%. Conversely, compared with the environmental samples (71.4%), the human-, animal-, and fermented food-derived samples presented higher proportions of chorismate production (M00022), ranging from 92.0% to 94.6% (Figure 2a). Certain KEGG modules were also present in some strains from human, animal, and fermented food sources and also contained certain KEGG modules (Figure 2b), including methionine degradation (M00035), lysine biosynthesis (M00031), and ornithine biosynthesis (M00763). These findings suggest that human-derived strains may be acquired from fermented foods and animals.

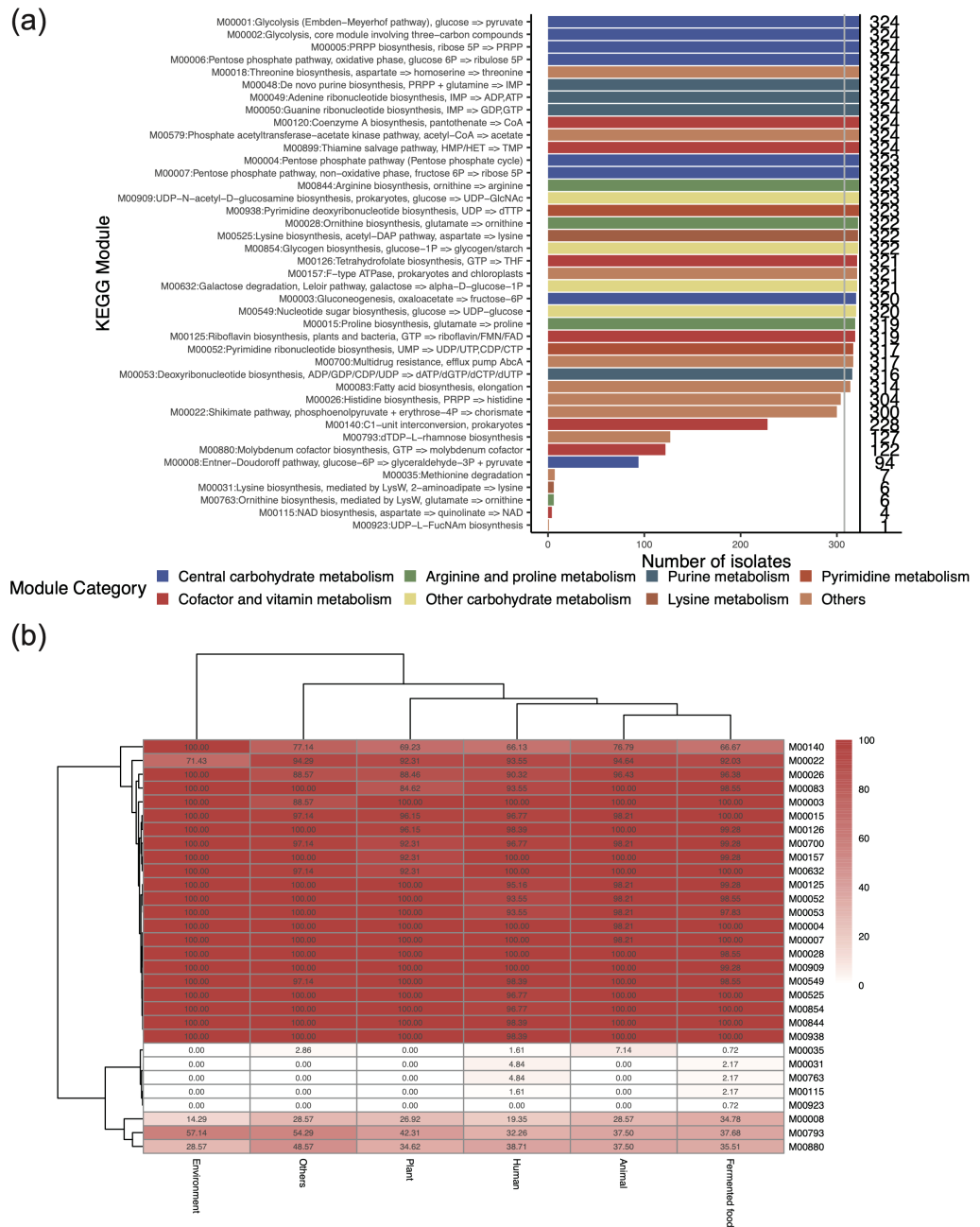


Figure 2. Presence of KEGG modules in the genomes of *L. plantarum*. **(a)** Prevalence of the KEGG modules among the strains. The different colors of each bar correspond to different KEGG categories. The gray vertical line indicates the 95% presence ratio. The numbers on the right represent the total number of strains within the correlated KEGG modules. **(b)** Prevalence of the KEGG modules from various hosts, excluding those modules present in all strains. The positive ratio of different modules in strains from different sources is represented by the numbers in each cell.

3.4. Profile of Glycoside Hydrolases (GHs) Reveals the Ability to Degrade a Variety of Carbohydrates

L. plantarum strains typically possess many genes encoding GHs, which help them to adapt to diverse environments. An annotation of the CAZymes revealed that *L. plantarum* can encode at least 28 different GH families, 11 of which are present in most strains ($\geq 95.0\%$) (Figure 3a). Notably, some GH families are not only prevalent in *L. plantarum* but also have multiple copies, which help to degrade carbohydrates (Figure 3b). For example, eight GH families were present in all strains: GH1 (β -glucosidases and β -galactosidases), GH2 (β -galactosidases, β -mannosidases, exo- β -

glucosaminidases), GH13 (α -glucosidases), GH25 (lysozymes), GH36 (α -galactosidases and α -N-acetylgalactosaminidases), GH65 (phosphorylases), GH73 (α -N-acetylglucosaminidases), and GH170 (6-phospho-N-acetylmeramidases). The average number of copies of the eight GH families in the 324 genomes ranged from 1.0 to 10.5. Interestingly, each genome contained an average of 10.5 copies of GH13 and 9.3 copies of GH1.

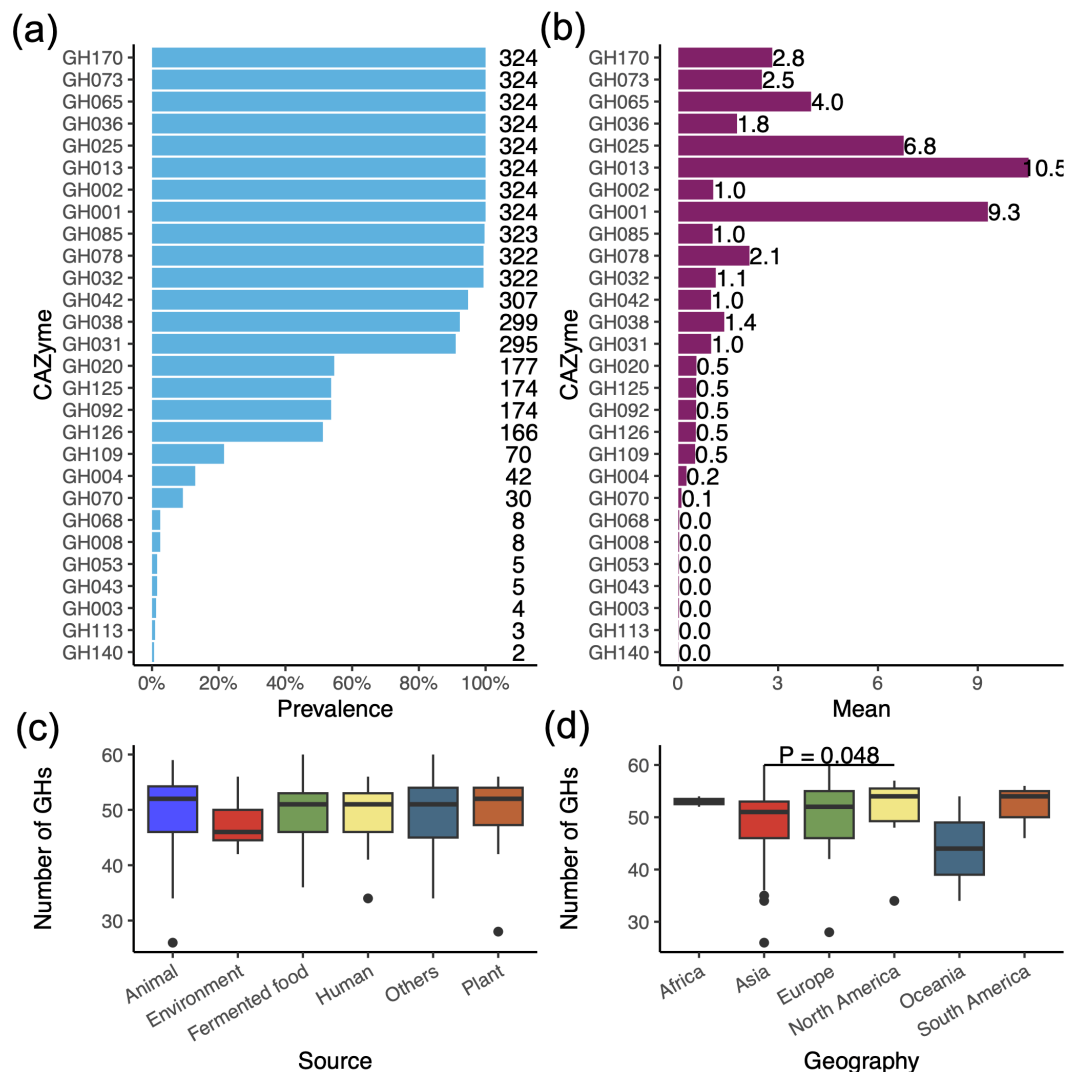


Figure 3. Glycoside Hydrolase (GH) family harbored in *L. plantarum* genomes. (a) Prevalence of GH families in the *L. plantarum* genomes. The numbers on the right indicate the total number of strains in correlated GH. (b) The average copy number of each GH family in the *L. plantarum* genomes. (c) The number of GHs in strains from different sources. (d) The number of GHs in strains from different continents.

For each genome, the average number of GH families across the 324 genomes was 49.5, with a range of 26 to 60. There was no significant difference among the strains from different sources. Additionally, the strains that are isolated from North America exhibited a higher number of GHs than those from Asia (Figure 3c,d).

3.5. Profile of Protease Types Reveals Their Ability to Degrade or Modify a Variety of Proteins

In microbes, peptidases are crucial enzymes that degrade proteins to acquire nutrients, process essential cellular components, and often play key roles in virulence, the stress response, and environmental adaptation. A total of 51 different protease families distributed

across seven protease types were detected in the *L. plantarum* genomes (Figure 4a). Metallo-, serine- and cysteine-peptidases were common members of the *L. plantarum* genome. Among these three peptidases, there were eight, nine, and six families found in the metallo-, serine-, and cysteine-type peptidases, respectively. These may be essential components of the *L. plantarum* genome. Moreover, the inhibitor I87 was detected in all *L. plantarum* genomes. Interestingly, the average number of copies of C26 per *L. plantarum* genome was 9.0 (Figure 4b), which is often involved in the turnover of folyl poly-gamma-glutamates.

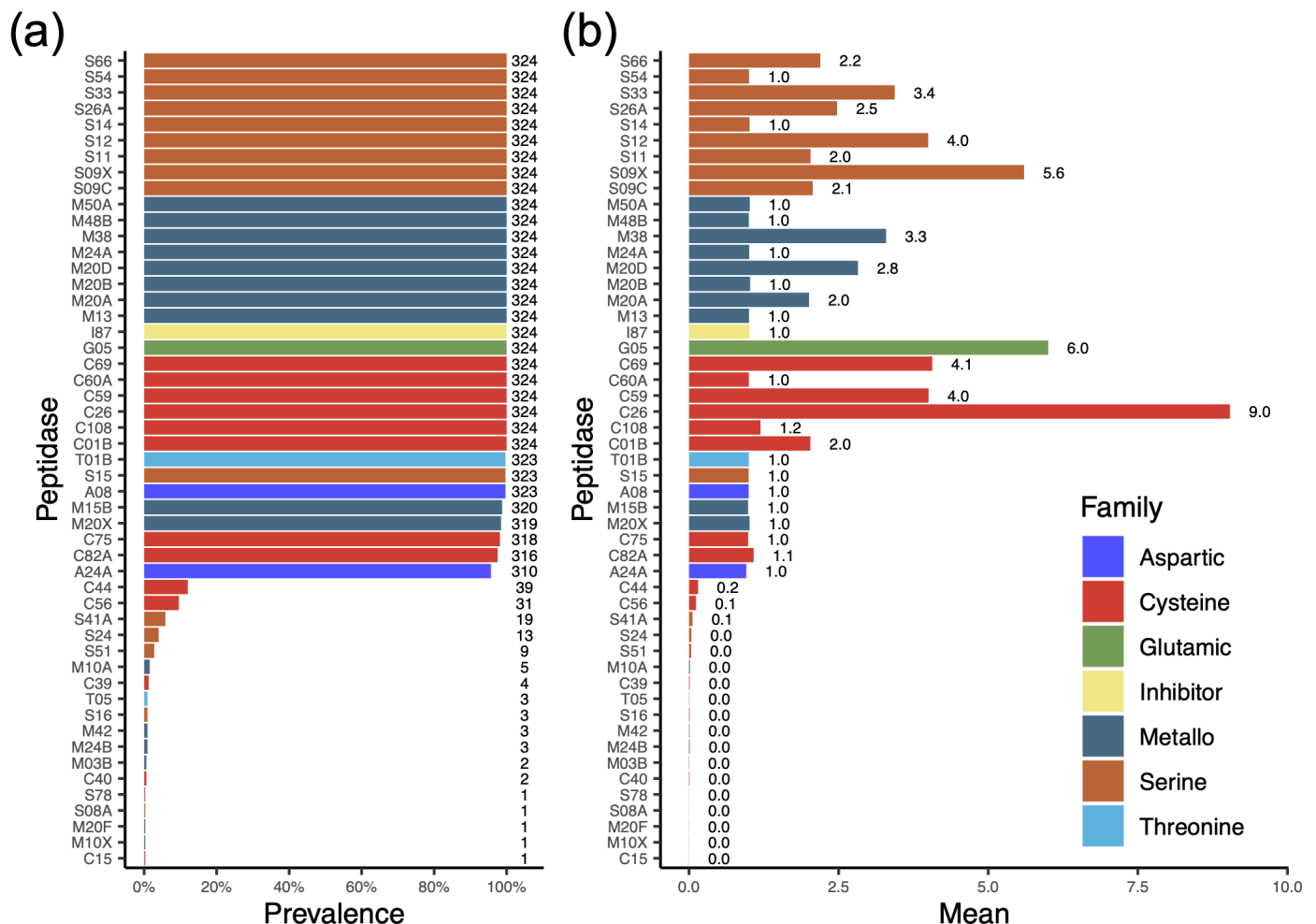


Figure 4. Protease families and inhibitors harbored in the *L. plantarum* genomes. (a) Prevalence of protease families and inhibitors in the *L. plantarum* genomes. The numbers on the right represent the total number of strains in correlated peptidase. (b) The average copy number of each protease family/inhibitor in the *L. plantarum* genomes.

3.6. The Antimicrobial Peptide (AMP) Facilitates the Antimicrobial Properties of *L. plantarum*

L. plantarum can produce AMPs that inhibit the growth of other bacteria. Therefore, we predicted which AMPs *L. plantarum* may harbor. AMP sequences were found in a total of 303 out of 324 genomes (93.5%, Figure 5a). On average, each genome from different origins contained an average number of 1.5 to 1.9 AMP sequences (Figure 5b). These AMPs belong to 77 different sequence types. The more common AMP sequence type was “VTGRLAVTLVGAPGPYVALIKTK”, which was found in 201 strains, and its mutant A18V, which was found in 62 strains (Figure 5c). To further verify the novelty of the 77 AMPs, we compared them against the known AMPs ($n = 44,406$). These results indicated that only one AMP sequence from *L. plantarum* SRCM100995 was identical to a partial

sequence (50/62) of an existing AMP (L12A07947) recorded in the LAMP2 database. The remaining sequences exhibited identities ranging from 23.3% to 96.0%. Notably, the most prevalent AMP ($n = 201$) demonstrated an identity of no more than 70.0% when compared to the known AMPs. These findings indicate that these AMPs are commonly carried by *L. plantarum*. Furthermore, we clustered these AMP sequences on the basis of a cut-off of 70.0% identity, forming 31 clusters. To further explore the sequence differences within the clusters, we constructed phylogenetic trees of AMPs from the same cluster and predicted their antimicrobial activity. The results revealed that, of the cluster represented by “VTGRLAVTLVGAPGPYVALIKTK” ($n = 201$), the variant A18V ($n = 62$) exhibited stronger antibacterial activity than the others (Figure 5d and Supplementary Table S3). In the cluster represented by “IATAILWAIKFMIVSFVGNVVKLIKPNRRYFGM” ($n = 31$), a number of 38 different AMP sequence types were found (Figure 5e). Among these mutations, the main mutation types were associated with a certain degree of decrease in antimicrobial activity.

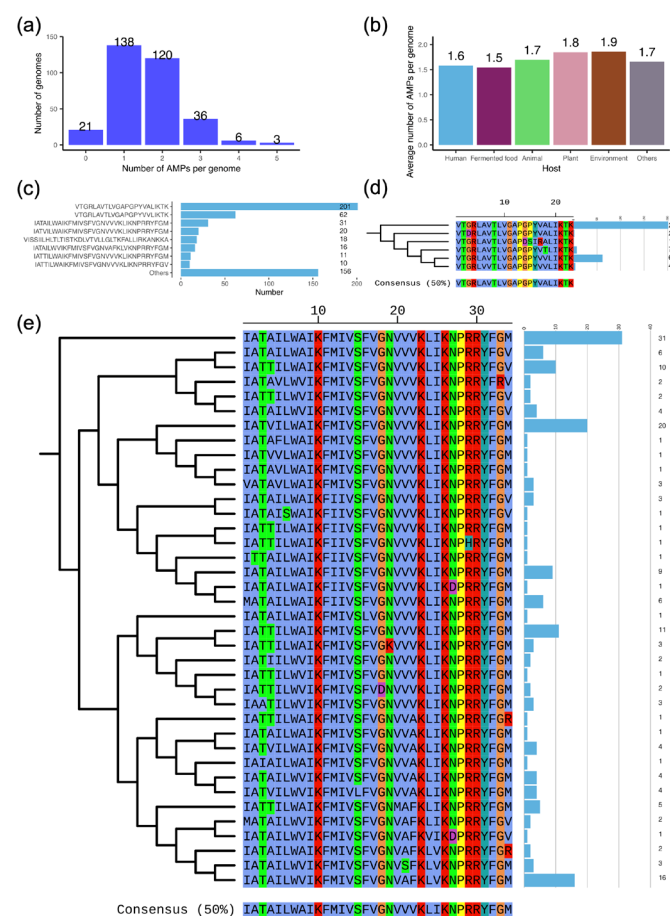


Figure 5. Antimicrobial peptides (AMPs) harbored in *L. plantarum* genomes. (a) Number of AMPs per genome. (b) Average number of AMPs per genome from various sources. (c) Number of the most common AMP sequences. (d–e) Phylogenetic tree of two clusters of AMP sequences. The sequence of each branch was labeled on the tree, and the number of each AMP was shown in the blue bar. The numbers on the right represent the total number of AMP sequences (c–e). Different amino acids were colored according to their chemical properties in (d–e).

3.7. Plasmids Mediate the Transmission of Antibiotic Resistance Genes Within the Lactobacillaceae Family

As probiotics, biosafety is an important aspect that should be considered. Therefore, we first scanned the genomes for virulence factors and found no positive hits. We then explored the ARGs harbored in the genome and identified a total of six ARGs. Two of

these (*vanH* and *vanY*) were found in almost all strains, except for those from the plant-related samples, of which the percentage of *vanH*-positive strains was 96.2% (Figure 6a). Another ARG, *vanT*, which confers vancomycin resistance, was present in 10.7% of the animal-derived strains and in 28.6% of the environmental strains. In addition to the vancomycin resistance genes, the *Tet(M)*, *ANT(6)-Ia*, and *mdeA* genes, which mediate tetracycline, aminoglycoside, and multidrug resistance, respectively, were also identified in these strains.

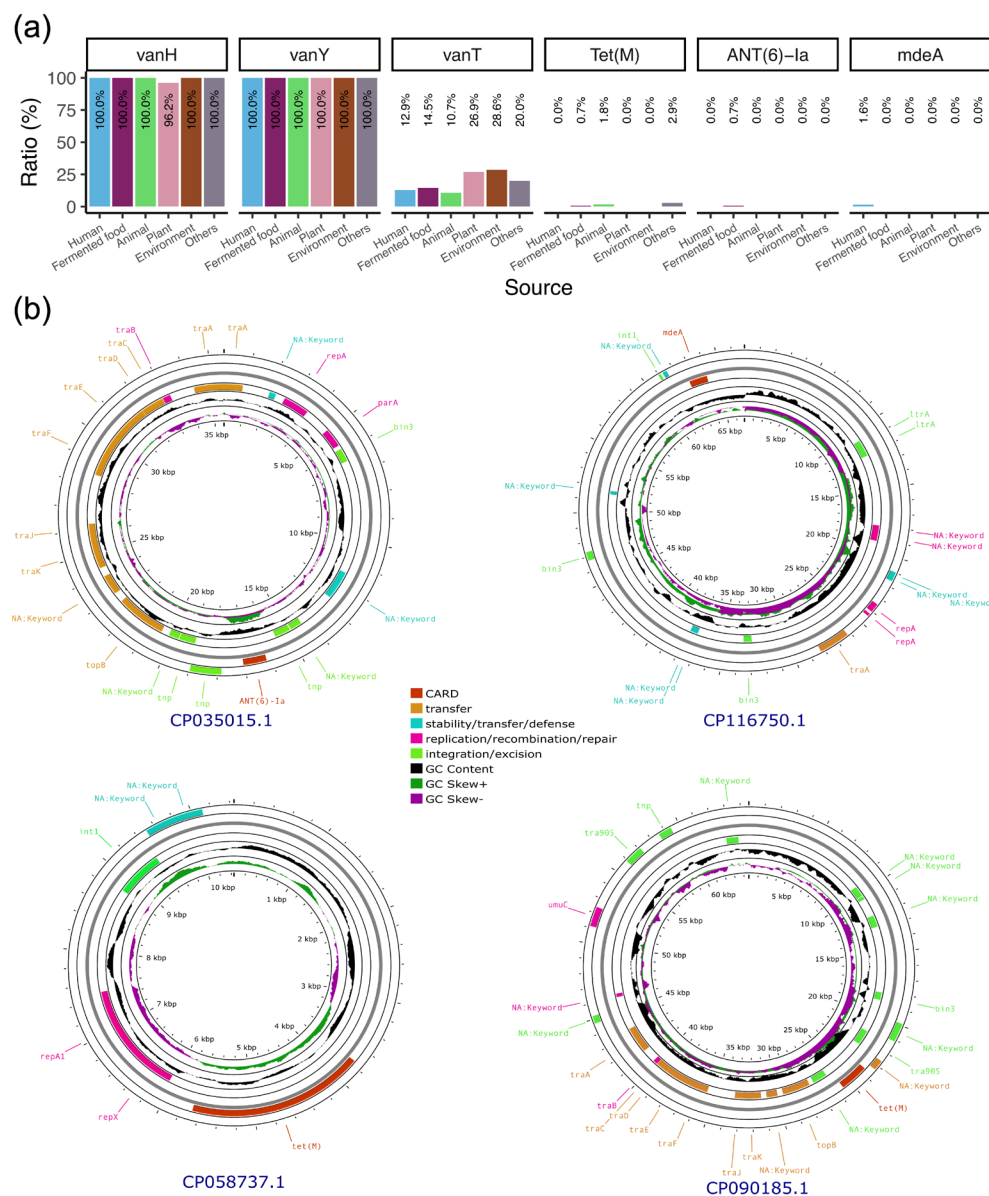


Figure 6. Antibiotic resistance genes (ARGs) that are carried by the *L. plantarum* strains and related plasmids. (a) Prevalence of ARGs in the strains from various sources. (b) The genetic structure of ARG-carrying plasmids. Note: The plasmids in *L. plantarum* A8 and *L. plantarum* J50 exhibited the same length with only two mismatches (6261/6263) and an ANI value of 0.990.

To determine whether these genes are mobile, we analyzed their associations with phages and plasmids. The results revealed that three plasmids carried the *Tet(M)* gene. Among these, the plasmid in *L. plantarum* ST could be transferred within the *Lactobacillaceae* family, as could the plasmid carrying the *ANT(6)-Ia* gene in *L. plantarum* 12_3. The plasmid carrying the *mdeA* gene in *L. plantarum* MWLp-12 could also potentially undergo transfer

(Figure 6b and Table 2). We also identified prophage sequences in the genomes, finding an average of one to seven fragments per genome belonging to the genus *Duplodnaviria* (Supplementary Table S4). Additionally, no ARGs were detected in the prophage fragments.

Table 2. Molecular characterization of the five ARG-carrying plasmids.

Accession	Strain Host	ARG	Replicon Type	Predicted Mobility	Predicted Host Range
CP035015	<i>L. plantarum</i> 12_3	<i>ANT(6)-Ia</i>	rep_cluster_707	Conjugative	<i>Lactobacillaceae</i>
CP058737	<i>L. plantarum</i> A8	<i>Tet(M)</i>	rep_cluster_2119	Non-mobilizable	<i>Lactobacillaceae</i>
CP090185	<i>L. plantarum</i> ST	<i>Tet(M)</i>	rep_cluster_167, rep_cluster_707	Conjugative	<i>Lactobacillaceae</i>
CP116750	<i>L. plantarum</i> MWLp-12	<i>mdeA</i>	rep_cluster_1328	Mobilizable	<i>Lactobacillales</i>
CP140093	<i>L. plantarum</i> J50	<i>Tet(M)</i>	rep_cluster_2119	Non-mobilizable	<i>Lactobacillaceae</i>

4. Discussion

In this study, we conducted a comparative genomic analysis of 324 complete *L. plantarum* genomes. The key findings of this study are as follows: (1) the strains were distributed unevenly from various origins; (2) *L. plantarum* may be able to produce some beneficial metabolites for hosts; (3) there was a high diversity of CAZymes and proteases in the *L. plantarum* genomes; (4) a large proportion of *L. plantarum* could produce AMPs as probiotics; and (5) the plasmid-mediated transmission of ARGs within the *Lactobacillaceae* family requires more attention in the future.

Our research revealed that *L. plantarum* has a total of 2403 core genome genes, which is higher than that of other *Lactobacillus* species, as reported [41,42]. These findings suggest that these genes are important for the survival of *L. plantarum*. Notably, among these core genes, 12.3% of these core genes have unknown functions, and their role in the survival of *L. plantarum* requires additional research, such as the use of the CRISPRi system [43].

The construction of phylogenetic trees can help us to better understand how different strains of the same species are transmitted from various sources. Previous evolutionary tree analyses of 54 *L. plantarum* strains did not reveal a close connection between the strains and their origins [12]. Our more extensive analysis similarly uncovered no inherent connections, indicating that *L. plantarum* sourced from diverse origins is disseminated via a multitude of pathways. For instance, plant-derived *L. plantarum* can be further transmitted to humans via the fermentation enrichment of fermented foods. However, the ability of colonization of the strains from fermented foods needs to be studied in the future.

Microorganisms in the gastrointestinal tract can synthesize essential amino acids de novo to maintain amino acid homeostasis in the host [44]. We found that *L. plantarum* can synthesize several essential amino acids for humans, including threonine, arginine, and lysine, which are crucial for maintaining the body's amino acid balance. Additionally, *L. plantarum* can also synthesize histidine, a conditionally essential amino acid that helps maintain nutritional balance in early infancy [45]. The supplementary of essential amino acids to hosts may bring benefits to hosts.

In addition to essential amino acids, *L. plantarum* can also synthesize various beneficial compounds. For example, it can produce IMP due to the presence of M00048. IMP enhances the flavor of fermented foods as a flavoring substance [46]. The genome of *L. plantarum* has also been found to metabolize and produce acetic acid, which inhibits the proliferation of pathogenic bacteria during fermentation and aids in enhancing the resistance of the host to pathogenic infections and regulating the immune system [47]. The presence of the thiamine salvage pathway results in the production of thiamine (vitamin B1). Thiamine is essential

for all living organisms because its active form, thiamine pyrophosphate, is an indispensable cofactor for enzymes that are involved in amino acid and carbohydrate metabolism [48].

An analysis of the CAZymes in *L. plantarum* indicates that it is a highly adaptable heterotroph that can utilize various complex carbohydrates as energy sources. For example, GH13 and GH1 are essential for the metabolism of complex carbohydrates [49], and multiple genes from these families increase the degradation ability. Additionally, the presence of various proteases not only helps *L. plantarum* to utilize different proteins effectively but also helps the host to digest and absorb proteins. These results suggest that *L. plantarum* has strong capabilities for utilizing polysaccharides and proteins, which endows it with significant environmental adaptability and assists the host in the digestion and absorption of nutrients.

As a probiotic, research has shown that its main mechanism of action is producing antimicrobial peptides. We found that *L. plantarum* primarily produces two types of antimicrobial peptides, represented by “VTGRLAVTLVGAPGPYVALIKTK” and “IATAIL-WAIFMIVSFVGNVVKLIKPNRRYFGM”. These two AMPs are not recorded in any antimicrobial peptide database. Furthermore, the antimicrobial activity of these mutants exhibits some degree of variability, which could impact strain selection during use.

We also analyzed the virulence factors and antibiotic resistance genes carried by the bacterium and found that our results are consistent with those of previous reports, indicating that there are no known virulence genes in its genome [9]. However, unlike previous studies, we discovered several genes that may undergo horizontal gene transfer. These genes are carried by plasmids that can be transferred within the *Lactobacillaceae* family.

The present study also has a limitation. While the in silico genomic analyses of *L. plantarum* offer insights into the genetic diversity, metabolic function, and antimicrobial properties of this species, this study lacks additional experimental validation. Specifically, the mobility of the plasmid-carrying ARGs, the biological activity of AMPs, and the metabolites require further investigation. These findings need to be verified to serve as a guide for future applications.

5. Conclusions

In this study, we conducted a comprehensive analysis of 324 *L. plantarum* genomes sourced from various origins and geographical locations. Our findings indicate transmissions occurring from foods to humans, other than from the environment. Furthermore, its ability to utilize a diverse range of carbohydrates and proteins supports its application in fermented foods. As a probiotic, one of the potential beneficial mechanisms of *L. plantarum* may be the production of various AMPs that inhibit other commensals, which can be considered as an intrinsic characteristic of this species. However, the plasmid-mediated dissemination of ARGs should be closely monitored within the fermented food industry.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes16080869/s1>, Table S1: Genomic characteristics of 324 strains of *L. plantarum*; Table S2: Number of Glycoside Hydrolases in *L. plantarum* genomes. Table S3: Prediction of minimum inhibitory concentration of AMPs against 34 type strains; Table S4: Prophage fragments in 324 *L. plantarum* genomes; Figure S1: Percentage of strains from different hosts (a) and locations (b) in each cluster in Figure 1. The numbers labeled above each bar represent the total number of strains in each cluster; Figure S2: Heatmap of the presence of KEGG modules in strains from various sources. The hierarchical clustering was based on Euclidean distance and adopted complete linkage. Both rows and columns were clustered.

Author Contributions: Conceptualization, C.B.; methodology, R.L.; software, R.L.; formal analysis, R.L.; writing—original draft preparation, R.L.; writing—review and editing, C.B.; visualization, R.L.; supervision, C.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Heilongjiang Provincial Key R&D Program, grant number 2024ZXDXB58.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The genomes included in this study can be found in <https://www.ncbi.nlm.nih.gov/datasets/genome/?taxon=1590>.

Acknowledgments: We would like to thank Yan Lin and Junyan Bie, from Beijing Yutai Biotechnology Co., Ltd., for their help in the data analysis of this study.

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

Abbreviations

The following abbreviations are used in this manuscript:

AMP	Antimicrobial Peptide
ANI	Average Nucleotide Identity
ARG	Antibiotic Resistance Gene
CARD	Comprehensive Antibiotic Resistance Database
CAZyme	Carbohydrate-Active Enzyme
GH	Glycoside Hydrolase
IMP	Inosine Monophosphate
KEGG	Kyoto Encyclopedia of Genes and Genomes
MGE	Mobile Genetic Element
MIC	Minimum Inhibitory Concentration
NCBI	National Center for Biotechnology Information

References

1. Siezen, R.J.; van Hylckama Vlieg, J.E.T. Genomic Diversity and Versatility of *Lactobacillus plantarum*, a Natural Metabolic Engineer. *Microb. Cell Fact.* **2011**, *10* (Suppl. S1), S3. [CrossRef] [PubMed]
2. Shah, A.B.; Baiseitova, A.; Zahoor, M.; Ahmad, I.; Ikram, M.; Bakhsh, A.; Shah, M.A.; Ali, I.; Idress, M.; Ullah, R.; et al. Probiotic Significance of *Lactobacillus* Strains: A Comprehensive Review on Health Impacts, Research Gaps, and Future Prospects. *Gut Microbes* **2024**, *16*, 2431643. [CrossRef] [PubMed]
3. Zheng, J.; Wittouck, S.; Salvetti, E.; Franz, C.M.A.P.; Harris, H.M.B.; Mattarelli, P.; O'Toole, P.W.; Pot, B.; Vandamme, P.; Walter, J.; et al. 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]
4. Chen, K.; Luan, X.; Liu, Q.; Wang, J.; Chang, X.; Snijders, A.M.; Mao, J.-H.; Secombe, J.; Dan, Z.; Chen, J.-H.; et al. Drosophila Histone Demethylase KDM5 Regulates Social Behavior through Immune Control and Gut Microbiota Maintenance. *Cell Host Microbe* **2019**, *25*, 537–552.e8. [CrossRef]
5. Cui, Y.; Wang, M.; Zheng, Y.; Miao, K.; Qu, X. The Carbohydrate Metabolism of *Lactiplantibacillus plantarum*. *Int. J. Mol. Sci.* **2021**, *22*, 13452. [CrossRef]
6. Tegopoulos, K.; Stergiou, O.S.; Kiouisi, D.E.; Tsifintaris, M.; Koletsou, E.; Papageorgiou, A.C.; Argyri, A.A.; Chorianopoulos, N.; Galanis, A.; Kolovos, P. Genomic and Phylogenetic Analysis of *Lactiplantibacillus plantarum* L125, and Evaluation of Its Anti-Proliferative and Cytotoxic Activity in Cancer Cells. *Biomedicines* **2021**, *9*, 1718. [CrossRef]
7. Huidrom, S.; Ngashangva, N.; Khumlianlal, J.; Sharma, K.C.; Mukherjee, P.K.; Devi, S.I. Genomic Insights from *Lactiplantibacillus plantarum* BRD3A Isolated from Atingba, a Traditional Fermented Rice-Based Beverage and Analysis of Its Potential for Probiotic and Antimicrobial Activity against Methicillin-Resistant *Staphylococcus Aureus*. *Front. Microbiol.* **2024**, *15*, 1357818. [CrossRef]
8. Tóth, A.G.; Judge, M.F.; Nagy, S.Á.; Papp, M.; Solymosi, N. A Survey on Antimicrobial Resistance Genes of Frequently Used Probiotic Bacteria, 1901 to 2022. *Eurosurveillance* **2023**, *28*, 2200272. [CrossRef]
9. Garcia-Gonzalez, N.; Bottacini, F.; van Sinderen, D.; Gahan, C.G.M.; Corsetti, A. Comparative Genomics of *Lactiplantibacillus plantarum*: Insights Into Probiotic Markers in Strains Isolated From the Human Gastrointestinal Tract and Fermented Foods. *Front. Microbiol.* **2022**, *13*, 854266. [CrossRef]

10. Evanovich, E.; de Souza Mendonça Mattos, P.J.; Guerreiro, J.F. Comparative Genomic Analysis of *Lactobacillus plantarum*: An Overview. *Int. J. Genom.* **2019**, *2019*, 4973214. [[CrossRef](#)]
11. Wu, M.; Tarrach, A.; Ghion, G.; Pakroo, S.; Giacomini, A.; Corich, V. A Critical Issue on Microbiological Cut-off Value of Ampicillin Resistance in *Lactiplantibacillus plantarum*. *J. Appl. Microbiol.* **2023**, *134*, lxad050. [[CrossRef](#)]
12. Martino, M.E.; Bayjanov, J.R.; Caffrey, B.E.; Wels, M.; Joncour, P.; Hughes, S.; Gillet, B.; Kleerebezem, M.; van Hijum, S.A.F.T.; Leulier, F. Nomadic Lifestyle of *Lactobacillus plantarum* Revealed by Comparative Genomics of 54 Strains Isolated from Different Habitats. *Environ. Microbiol.* **2016**, *18*, 4974–4989. [[CrossRef](#)]
13. Chaumeil, P.A.; Mussig, A.J.; Hugenholtz, P.; Parks, D.H. GTDB-Tk: A Toolkit to Classify Genomes with the Genome Taxonomy Database. *Bioinformatics* **2020**, *36*, 1925–1927. [[CrossRef](#)]
14. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: Assessing the Quality of Microbial Genomes Recovered from Isolates, Single Cells, and Metagenomes. *Genome Res.* **2015**, *25*, 1043–1055. [[CrossRef](#)] [[PubMed](#)]
15. Pritchard, L.; Glover, R.H.; Humphris, S.; Elphinstone, J.G.; Toth, I.K. Genomics and Taxonomy in Diagnostics for Food Security: Soft-Rotting Enterobacterial Plant Pathogens. *Anal. Methods* **2016**, *8*, 12–24. [[CrossRef](#)]
16. Seemann, T. Prokka: Rapid Prokaryotic Genome Annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [[CrossRef](#)] [[PubMed](#)]
17. Tonkin-Hill, G.; MacAlasdair, N.; Ruis, C.; Weimann, A.; Horesch, G.; Lees, J.A.; Gladstone, R.A.; Lo, S.; Beaudoin, C.; Floto, R.A.; et al. Producing Polished Prokaryotic Pangenomes with the Panaroo Pipeline. *Genome Biol.* **2020**, *21*, 180. [[CrossRef](#)]
18. Huerta-Cepas, J.; Szklarczyk, D.; Heller, D.; Hernández-Plaza, A.; Forslund, S.K.; Cook, H.; Mende, D.R.; Letunic, I.; Rattei, T.; Jensen, L.J.; et al. EggNOG 5.0: A Hierarchical, Functionally and Phylogenetically Annotated Orthology Resource Based on 5090 Organisms and 2502 Viruses. *Nucleic Acids Res.* **2019**, *47*, D309–D314. [[CrossRef](#)]
19. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; Von Haeseler, A.; Lanfear, R.; Teeling, E. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **2020**, *37*, 1530–1534. [[CrossRef](#)]
20. Tonkin-Hill, G.; Lees, J.A.; Bentley, S.D.; Frost, S.D.W.; Corander, J. Fast Hierarchical Bayesian Analysis of Population Structure. *Nucleic Acids Res.* **2019**, *47*, 5539–5549. [[CrossRef](#)]
21. Zhou, Z.; Tran, P.Q.; Breister, A.M.; Liu, Y.; Kieft, K.; Cowley, E.S.; Karaoz, U.; Anantharaman, K. METABOLIC: High-Throughput Profiling of Microbial Genomes for Functional Traits, Metabolism, Biogeochemistry, and Community-Scale Functional Networks. *Microbiome* **2022**, *10*, 33. [[CrossRef](#)] [[PubMed](#)]
22. Zhou, S.; Liu, B.; Zheng, D.; Chen, L.; Yang, J. VFDB 2025: An Integrated Resource for Exploring Anti-Virulence Compounds. *Nucleic Acids Res.* **2025**, *53*, D871–D877. [[CrossRef](#)] [[PubMed](#)]
23. Alcock, B.P.; Huynh, W.; Chalil, R.; Smith, K.W.; Raphenya, A.R.; Wlodarski, M.A.; Edalatmand, A.; Petkau, A.; Syed, S.A.; Tsang, K.K.; et al. CARD 2023: Expanded Curation, Support for Machine Learning, and Resistome Prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res.* **2023**, *51*, D690–D699. [[CrossRef](#)]
24. Camargo, A.P.; Roux, S.; Schulz, F.; Babinski, M.; Xu, Y.; Hu, B.; Chain, P.S.G.; Nayfach, S.; Kyrpides, N.C. Identification of Mobile Genetic Elements with geNomad. *Nat. Biotechnol.* **2023**, *42*, 1303–1312. [[CrossRef](#)]
25. Zheng, K.; Sun, J.; Liang, Y.; Kong, L.; Paez-Espino, D.; Mcminn, A.; Wang, M. VITAP: A High Precision Tool for DNA and RNA Viral Classification Based on Meta-Omic Data. *Nat. Commun.* **2025**, *16*, 2226. [[CrossRef](#)]
26. Robertson, J.; Bessonov, K.; Schonfeld, J.; Nash, J.H.E. Universal Whole-Sequence-Based Plasmid Typing and Its Utility to Prediction of Host Range and Epidemiological Surveillance. *Microb. Genom.* **2020**, *6*, mgen000435. [[CrossRef](#)]
27. Santos-Júnior, C.D.; Pan, S.; Zhao, X.-M.; Coelho, L.P. Macrel: Antimicrobial Peptide Screening in Genomes and Metagenomes. *PeerJ* **2020**, *8*, e10555. [[CrossRef](#)]
28. Mondal, R.K.; Sen, D.; Arya, A.; Samanta, S.K. Developing Anti-Microbial Peptide Database Version 1 to Provide Comprehensive and Exhaustive Resource of Manually Curated AMPs. *Sci. Rep.* **2023**, *13*, 17843. [[CrossRef](#)]
29. Wang, G.; Li, X.; Wang, Z. APD3: The Antimicrobial Peptide Database as a Tool for Research and Education. *Nucleic Acids Res.* **2016**, *44*, D1087–D1093. [[CrossRef](#)]
30. Pirtskhalava, M.; Armstrong, A.A.; Grigolava, M.; Chubinidze, M.; Alimbarashvili, E.; Vishnepolsky, B.; Gabrielian, A.; Rosenthal, A.; Hurt, D.E.; Tartakovsky, M. DBAASP v3: Database of Antimicrobial/Cytotoxic Activity and Structure of Peptides as a Resource for Development of New Therapeutics. *Nucleic Acids Res.* **2021**, *49*, D288–D297. [[CrossRef](#)]
31. Shi, G.; Kang, X.; Dong, F.; Liu, Y.; Zhu, N.; Hu, Y.; Xu, H.; Lao, X.; Zheng, H. DRAMP 3.0: An Enhanced Comprehensive Data Repository of Antimicrobial Peptides. *Nucleic Acids Res.* **2022**, *50*, D488–D496. [[CrossRef](#)]
32. Ye, G.; Wu, H.; Huang, J.; Wang, W.; Ge, K.; Li, G.; Zhong, J.; Huang, Q. LAMP2: A Major Update of the Database Linking Antimicrobial Peptides. *Database* **2020**, *2020*, baaa061. [[CrossRef](#)]
33. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
34. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for Clustering the next-Generation Sequencing Data. *Bioinformatics* **2012**, *28*, 3150–3152. [[CrossRef](#)] [[PubMed](#)]

35. Edgar, R.C. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)] [[PubMed](#)]
36. Wan, F.; Torres, M.D.T.; Peng, J.; de la Fuente-Nunez, C. Deep-Learning-Enabled Antibiotic Discovery through Molecular de-Extinction. *Nat. Biomed. Eng.* **2024**, *8*, 854–871. [[CrossRef](#)] [[PubMed](#)]
37. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v6: Recent Updates to the Phylogenetic Tree Display and Annotation Tool. *Nucleic Acids Res.* **2024**, *52*, W78–W82. [[CrossRef](#)]
38. Grant, J.R.; Enns, E.; Marinier, E.; Mandal, A.; Herman, E.K.; Chen, C.-Y.; Graham, M.; Van Domselaar, G.; Stothard, P. Proksee: In-Depth Characterization and Visualization of Bacterial Genomes. *Nucleic Acids Res.* **2023**, *51*, W484–W492. [[CrossRef](#)]
39. Kolde, R.; Kolde, M.R. Maintainer Raivo Package ‘pheatmap’. *Bioconductor* **2015**, *1*, 790.
40. Brunson, J.C. Ggalluvial: Alluvial Plots in “Ggplot2”. *ggalluvial* **2020**, *5*, 2017.
41. Zhang, L.; Kulyar, M.F.; Niu, T.; Yang, S.; Chen, W. Comparative Genomics of *Limosilactobacillus Reuteri* YLR001 Reveals Genetic Diversity and Probiotic Properties. *Microorganisms* **2024**, *12*, 1636. [[CrossRef](#)]
42. You, L.; Lv, R.; Jin, H.; Ma, T.; Zhao, Z.; Kwok, L.-Y.; Sun, Z. A Large-Scale Comparative Genomics Study Reveals Niche-Driven and within-Sample Intra-Species Functional Diversification in *Lactocaseibacillus rhamnosus*. *Food Res. Int.* **2023**, *173*, 113446. [[CrossRef](#)] [[PubMed](#)]
43. Myrbråten, I.S.; Wiull, K.; Salehian, Z.; Håvarstein, L.S.; Straume, D.; Mathiesen, G.; Kjos, M. CRISPR Interference for Rapid Knockdown of Essential Cell Cycle Genes in *Lactobacillus plantarum*. *mSphere* **2019**, *4*, e00007–e00019. [[CrossRef](#)] [[PubMed](#)]
44. Lin, R.; Liu, W.; Piao, M.; Zhu, H. A Review of the Relationship between the Gut Microbiota and Amino Acid Metabolism. *Amino Acids* **2017**, *49*, 2083–2090. [[CrossRef](#)] [[PubMed](#)]
45. Man, L.-L.; Xiang, D.-J. Effect of LuxS/AI-2-Mediated Quorum Sensing System on Bacteriocin Production of *Lactobacillus plantarum* NMD-17. *Folia Microbiol.* **2023**, *68*, 855–866. [[CrossRef](#)]
46. Hu, Q.; Qi, Y.; Liu, C.; Chen, Q.; Mai, X.; Zhu, Z.; Ma, B. Effects of *Lactobacillus plantarum* Cofermentation on the Flavor and Taste Characteristics of Mango Vinegar. *J. Food Meas. Charact.* **2024**, *18*, 3744–3756. [[CrossRef](#)]
47. Mann, E.R.; Lam, Y.K.; Uhlig, H.H. Short-Chain Fatty Acids: Linking Diet, the Microbiome and Immunity. *Nat. Rev. Immunol.* **2024**, *24*, 577–595. [[CrossRef](#)]
48. Gonçalves, C.; Gonçalves, P. Multilayered Horizontal Operon Transfers from Bacteria Reconstruct a Thiamine Salvage Pathway in Yeasts. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 22219–22228. [[CrossRef](#)]
49. Iarusso, I.; Mahony, J.; Pannella, G.; Lombardi, S.J.; Gagliardi, R.; Coppola, F.; Pellegrini, M.; Succi, M.; Tremonte, P. Diversity of *Lactiplantibacillus plantarum* in Wild Fermented Food Niches. *Foods* **2025**, *14*, 1765. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.