



## OPEN ACCESS

## EDITED BY

Waseem Safdar,  
National University of Medical Sciences  
(NUMS), Pakistan

## REVIEWED BY

Ahmed A. Zaky,  
National Research Centre, Egypt  
Osman Erkmen,  
Istanbul Arel University, Türkiye  
Ayla Arslaner,  
Bayburt University, Türkiye

## \*CORRESPONDENCE

Yang Zhennai  
✉ yangzhennai@163.com

RECEIVED 29 April 2023

ACCEPTED 18 May 2023

PUBLISHED 14 June 2023

## CITATION

Aziz T, Naveed M, Shabbir MA, Sarwar A, Ali Khan A, Zhennai Y, Alharbi M, Alsahammari A and Alasmari AF (2023) Comparative genomics of food-derived probiotic *Lactiplantibacillus plantarum* K25 reveals its hidden potential, compactness, and efficiency.  
*Front. Microbiol.* 14:1214478.  
doi: 10.3389/fmicb.2023.1214478

## COPYRIGHT

© 2023 Aziz, Naveed, Shabbir, Sarwar, Ali Khan, Zhennai, Alharbi, Alsahammari and Alasmari. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Comparative genomics of food-derived probiotic *Lactiplantibacillus plantarum* K25 reveals its hidden potential, compactness, and efficiency

Tariq Aziz<sup>1,2</sup>, Muhammad Naveed<sup>3</sup>, Muhammad Aqib Shabbir<sup>3</sup>, Abid Sarwar<sup>1</sup>, Ayaz Ali Khan<sup>4</sup>, Yang Zhennai<sup>1\*</sup>, Metab Alharbi<sup>5</sup>, Abdulrahman Alsahammari<sup>5</sup> and Abdullah F. Alasmari<sup>5</sup>

<sup>1</sup>Beijing Advanced Innovation Center for Food Nutrition and Human Health, Beijing Engineering and Technology Research Center for Food Additives, Beijing Technology and Business University, Beijing, China, <sup>2</sup>Laboratory of Animal Health, Food Hygiene, and Quality, Department of Agriculture, University of Ioannina, Arta, Greece, <sup>3</sup>Department of Biotechnology, Faculty of Science and Technology, University of Central Punjab, Lahore, Pakistan, <sup>4</sup>Department of Biotechnology, University of Malakand, Chakdara, Pakistan, <sup>5</sup>Department of Pharmacology and Toxicology, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia

This study aimed to investigate the intricate genetic makeup of the *Lactiplantibacillus plantarum* K25 strain by conducting a comprehensive analysis of comparative genomics. The results of our study demonstrate that the genome exhibits a high-level efficiency and compactness, comprising a total of 3,199 genes that encode proteins and a GC content of 43.38%. The present study elucidates the evolutionary lineage of *Lactiplantibacillus plantarum* strains through an analysis of the degree of gene order conservation and synteny across a range of strains, thereby underscoring their closely interrelated evolutionary trajectories. The identification of various genetic components in the K25 strain, such as bacteriocin gene clusters and prophage regions, highlights its potential utility in diverse domains, such as biotechnology and medicine. The distinctive genetic elements possess the potential to unveil innovative therapeutic and biotechnological remedies in future. This study provides a comprehensive analysis of the *L. plantarum* K25 strain, revealing its remarkable genomic potential and presenting novel prospects for utilizing its unique genetic features in diverse scientific fields. The present study contributes to the existing literature on *Lactiplantibacillus plantarum* and sets the stage for prospective investigations and practical implementations that leverage the exceptional genetic characteristics of this adap organism.

## KEYWORDS

*L. plantarum* K25, comparative genomics, bacteriocin, nucleotide identity, genomic potential

## Introduction

Lactic acid bacteria (LAB) are generally regarded as safe (GRAS) because of their diverse beneficial properties in the food and dairy industries (Sarwar et al., 2018; Aziz et al., 2023). Among LAB, lactobacilli are the most widespread (Min et al., 2020; Aziz et al., 2022a,b). *Lactiplantibacillus* K25 is a lactic acid bacterium strain that is widespread in fermented foods and drinks. It is a

probiotic strain that has garnered a lot of interest because of its potential health advantages, specifically, its capacity to promote gut health and strengthen the immune system (Aziz et al., 2022b). *L. plantarum* K25 has been the focus of multiple investigations in recent years, with research indicating that it may have a variety of uses ranging from treating inflammatory bowel illness to avoiding some forms of cancer. The strain has also been used as a starting culture in the fermentation of foods such as kimchi and sauerkraut. Despite its potential benefits, there is still a lot more to learn about *L. plantarum* K25, and more studies are needed to properly understand it (Garcia-Gonzalez et al., 2021).

An *L. plantarum* K25 comparative genomics study contributes to the expanding body of research on the genetic makeup and possible advantages of this probiotic strain. Previous research has shown that *L. plantarum* K25 can enhance gut health, stimulate the immune system, and perhaps prevent certain illnesses (An et al., 2023). A comparative genomics study adds to earlier studies by elucidating the genetic adaptations that have allowed *L. plantarum* K25 to survive in specific ecological niches. The research also discovers novel genes and metabolic pathways in *Lactiplantibacillus* K25 that might have uses in probiotics and biotechnology (Barbosa et al., 2021).

Although comparative genomics studies on *Lactiplantibacillus* have shed light on the genetic makeup of this probiotic strain, there is still a knowledge gap in our understanding of its genome. For example, one study did not address the role of certain genes in the metabolism of *L. plantarum* K25 nor did it investigate the mechanisms underlying its probiotic properties (Aziz et al., 2022a). Additionally, the study did not explore the potential impact of environmental factors on the genome of *Lactiplantibacillus*. Further research is needed to fill these knowledge gaps and deepen our understanding of the genetic adaptations of this strain, as well as its potential applications in biotechnology and probiotics (Tegopoulos et al., 2021).

In this study, a range of methodologies were used to analyze the genome of *L. plantarum* K25. The first step was data collection, followed by quality assessment, gene annotation, and average nucleotide identity. The pan-genome was then estimated and modeled, and phylogenetic analysis and population structure analysis were carried out. The genomic plasticity was also analyzed. Functional annotation and enrichment analyses as well as taxonomy, phylogenomic, and evolutionary analyses were performed. Mobile elements, insertion sequences, bacteriocins, and CRISPR-Cas systems were predicted, and an essentiality analysis was conducted. Finally, genes related to probiotic features were identified. Overall, these methodologies provided important insights into the genetic makeup of *L. plantarum* K25 and its potential uses in biotechnology and probiotics.

In conclusion, the comparative genomics study on *L. plantarum* K25 has provided valuable insights into the genetic makeup of this probiotic strain. This study identified unique genes and metabolic pathways in *L. plantarum* K25 that may have potential applications in the fields of probiotics and biotechnology. However, there are still knowledge gaps in our understanding of the genome of *L. plantarum* K25, particularly in the role of certain genes and the mechanisms underlying its probiotic properties (Zhao et al., 2022). Further research is needed to fill

these gaps and fully unlock the potential of *Lactiplantibacillus* K25 for improving human health and advancing biotechnological research. Overall, this study highlights the importance of genomics research in uncovering the hidden potential of microorganisms and advancing our understanding of the natural world (Jeong et al., 2021).

## Materials and methods

### Bacterial strain and culture condition

In 2015, *Lactiplantibacillus plantarum* K25 (*L. plantarum* K25) was obtained from Tibetan Kefir grains and subsequently preserved as frozen stocks at a temperature of  $-9^{\circ}\text{C}$  in MRS broth that was supplemented with 20% (v/v) glycerol. The identification of these strains was primarily based on Gram staining, catalase testing, and cellular morphology. The identification of strains was carried out using the API 50 CHL test (bioMérieux, Marcy-l'Étoile, France), and 16S rDNA sequencing analysis was carried out as previously described (Wang et al., 2015; Zhang et al., 2017, 2020; Yunyun et al., 2018; Aziz et al., 2020a,b,c, 2021, 2022a,b).

### DNA extraction and whole-genome sequencing

The Wizard<sup>®</sup> Genomic DNA Purification Kit from Promega, Madison, WI, USA was utilized to extract the genomic DNA, which was subsequently quantified using a TBS-380 fluorometer (Turner Bio Systems Inc., Sunnyvale, CA, USA). The insert size for the quantification was 15 kb. For further analysis, DNA of superior quality was utilized, with an OD<sub>260/280</sub> ratio ranging from 1.8 to 2.0 and a quantity exceeding 20  $\mu\text{g}$ . The sheared fragments were utilized to prepare Illumina sequencing libraries using the NEXTflex<sup>™</sup> Rapid DNA-Seq Kit. In brief, a total of 50 prime ends underwent initial end-repair and phosphorylation. Subsequently, 30 termini were subjected to A-tailing and ligation with sequencing adapters. The third phase involved the amplification of the adapters-ligated products through PCR for enrichment purposes. The libraries that had been prepared were subjected to paired-end Illumina sequencing, with a read length of 150 bp for each end, utilizing an Illumina HiSeq X Ten platform. The genome of the chosen strain, *L. plantarum* K25, was sequenced using both single molecule real-time (SMRT) technology and Illumina sequencing platforms. The resulting sequence was assigned the accession number GCA\_003020005.1. The complexity of the genome was assessed utilizing the Illumina data (Churro et al., 2012). The genomic DNA was isolated in accordance with the manufacturer's protocol, utilizing the Qiagen DNA extraction kit. The isolation process was carried out from March to May 2019.

### Data collection

On April 14, 2023, a customized bash script was used to retrieve 10 genomes of the *Lactiplantibacillus plantarum* that were

accessible from the NCBI RefSeq database (<https://www.ncbi.nlm.nih.gov>). Our sequenced genomes were the *Lactiplantibacillus plantarum* K25 strain that was earlier submitted to NCBI with the specifically allocated accession number GCA\_003020005.1. To make it simple to distinguish between various genomes, the genomes were changed from assembly names to strain names.

## Gene annotation and average nucleotide identity

MetaGeneMark (available at <http://opal.biology.gatech.edu/GeneMark>), was utilized to perform gene coordinates prediction, protein sequence, and nucleotide sequence prediction (Gemayel et al., 2022). Additionally, Prokka v1.12 was also employed to develop pipelines for tRNA, tmRNA, rRNA, and CRISPR sequence prediction using Aragorn (available at <http://www.ansikte.se/ARAGORN>), and the genome sequence was submitted in the FASTA DNA file to predict the tRNA, tmRNA, and CRISPR sequences present in the genome (Gemayel et al., 2022).

To investigate the taxonomic boundaries of the genomes, average nucleotide identity (ANI) was determined using a genome-to-genome distance calculator (GGDC) (accessible at <http://ggdc.dsmz.de>) that uses BLAST + as the similarity search program. All of the 10 retrieved genomes along with our reference genome were submitted to the server for analysis (Auch et al., 2010).

## Genome-based taxonomy analysis

For the analysis of genome-based taxonomy, the Type (Strain) Genome Server (TYGS) was utilized, which is freely accessible (<https://tygs.dsmz.de>). The whole-genome sequence was submitted in the FASTA DNA file for analysis, and the results were further analyzed to evaluate the genome-based taxonomy (Meier-Kolthoff and Göker, 2019).

## Pan-genome estimation

The pan-genome analysis was performed using the PanX: Pan-Genome Analysis and Exploration online server that is freely available (<https://pangenome.org>). The species was selected as *Lactiplantibacillus plantarum* for the analysis, and the core genes, distribution ranking, and strain trees were obtained (Ding et al., 2018).

## Genomic plasticity analysis

The analysis of gene plasticity is a valuable tool in identifying drug targets and developing innovative strategies to combat bacterial infections. Understanding the evolution of bacterial genomes and how they adapt to new environments is crucial in this process. By identifying the genomic regions involved in gene plasticity, we can gain insights into the diversity of bacterial populations, the distribution of virulence factors, and the spread

TABLE 1 Retrieved genomic sequences of *Lactiplantibacillus plantarum* from NCBI.

Sr	Strain	Accession no.
1	<i>Lactiplantibacillus plantarum</i> strain 5-2	GCA_001278015.1
2	<i>Lactiplantibacillus plantarum</i> strain HFC8	GCA_001302645.1
3	<i>Lactiplantibacillus plantarum</i> strain LZ95	GCA_001484005.1
4	<i>Lactiplantibacillus plantarum</i> strain JBE245	GCA_001596095.1
5	<i>Lactiplantibacillus plantarum</i> strain ZS2058	GCF_001296095.1
6	<i>Lactiplantibacillus plantarum</i> strain WLPL04	GCF_001331925.2
7	<i>Lactiplantibacillus plantarum</i> strain Zhang-LL	GCF_001581895.1
8	<i>Lactiplantibacillus plantarum</i> strain CAUH2	GCF_001617525.2
9	<i>Lactiplantibacillus plantarum</i> strain SRCM100442	GCF_0009913655.1
10	<i>Lactiplantibacillus plantarum</i> strain K25	GCF_003020005.1

of antibiotic-resistance genes. The genomic plasticity analysis was performed by IslandViewer 4 (freely accessible at <https://www.pathogenomics.sfu.ca>). The genome of *L. plantarum* K25 was selected for analysis, and the results were interpreted (Dhillon et al., 2013).

## Prediction of prophage regions, probiotic potential, and bacteriocins

The prediction of prophage regions in the genome of the *L. plantarum* K25 strain was performed using the PHAge Search Tool Enhanced Release (PHASTER), which is freely accessible (<https://phaster.ca>). The prophage start-end regions, CDS, and percentage GC were predicted as the output that was further interpreted (Arndt et al., 2016). The identification of Bacteriocins in the K25 genome was analyzed using the BAGEL4 online server (freely accessible at <http://bagel4.molgenrug.nl>). The FASTA DNA file of the genome was given as the input, and the results were interpreted to determine the bacteriocin-producing genes in the genome under study (van Heel et al., 2018).

## Results

### Data collection and sorting

The retrieved genomic data were saved in the FASTA format for further analysis. There were a total of 10 genomes retrieved from the NCBI RefSeq database. The retrieved reference genomes are shown in Table 1.

### Gene annotation and average nucleotide identity

For the microbial genome of *L. plantarum* K25, MetaGeneMark was used to predict genes. A total of 3,199 genes with an average length of 800–1,000 bp were predicted. With an average intergenic

TABLE 2 tRNA codon and tRNA anticodon frequency.

tRNA anticodon frequency			
AAA Phe	AGA Ser	ATA Tyr	ACA Cys
GAA Phe 2	GGA Ser 1	GTA Tyr 2	GCA Cys 1
TAA Leu 2	TGA Ser 1	TTA Stop	TCA Sec 1
CAA Leu 1	CGA Ser 1	CTA Pyl	CCA Trp 1
AAG Leu 1	AGG Pro	ATG His	ACG Arg 2
GAG Leu	GGG Pro	GTG His 2	GCG Arg
TAG Leu 1	TGG Pro 2	TTG Gln 2	TCG Arg
CAG Leu 1	CGG Pro 1	CTG Gln 1	CCG Arg 1
AAT Ile	AGT Thr 1	ATT Asn	ACT Ser
GAT Ile 3	GGT Thr 1	GTT Asn 5	GCT Ser 1
TAT Ile	TGT Thr 1	TTT Lys 2	TCT Arg 1
CAT Met 5	CGT Thr 1	CTT Lys 1	CCT Arg 1
AAC Val	AGC Ala	ATC Asp	ACC Gly
GAC Val	GGC Ala	GTC Asp 3	GCC Gly 2
TAC Val 3	TGC Ala 2	TTC Glu 2	TCC Gly 2
CAC Val	CGC Ala 1	CTC Glu 1	CCC Gly 1
Ambiguous: 2			

distance of 1,000 bp, the predicted genes were dispersed throughout the genome.

Various metrics, including sensitivity, specificity, and positive predictive value (PPV), were used to assess the gene prediction accuracy. The number of genuine genes that were accurately identified was 3,120, which represents the sensitivity of the gene predictions. The number of anticipated genes that were true positives, or specificity, was 1,559 throughout the genome.

Our analysis revealed a wide range of tRNA, tmRNA, and rRNA sequences in the genomes of the organisms studied. For tRNA, we found a variety of types, including tRNA-Val, tRNA-Met, tRNA-Ala, and others. The predicted secondary structures of the tRNA molecules were also analyzed, and we observed variations in the anticodon loops and stems.

In addition to tRNA, we also predicted the presence of tmRNA in the genomes analyzed. tmRNA is involved in the degradation of aberrant proteins during translation, and its prediction can provide insight into protein quality control mechanisms. The total number of tRNA genes was observed to be 69. We found a diverse range of tmRNA sequences, with variations in the size, structure, and location within the genome.

tRNA codon and tRNA anticodon frequencies are shown in Table 2.

Our analysis revealed a wide range of ANI values among the genomes studied. ANI values ranged from 77.4 to 84.8%, indicating significant differences in genomic similarity among the organisms. We observed that organisms within the same genus had higher ANI values, indicating a higher degree of genomic similarity, while organisms from different genera had lower ANI values, indicating a lower degree of genomic similarity.

We also observed that ANI values varied within the same species. In some cases, ANI values were close to 80%, indicating a high degree of genomic similarity, while in other cases, ANI values were lower, indicating a lower degree of genomic similarity. This variability may be due to differences in genomic content or evolution within the same species. Table 3 shows the detailed results of the ANI analysis.

## Pan-genome estimation and modeling

According to our findings, the pan-genome consists of thousands of genes and is quite big and diversified. We found that there are various types of genomes, one consisting of genes shared by all creatures and the other consisting of genes shared by only a subset of those organisms. Different prokaryotes had different-sized auxiliary genomes, which we interpret to mean they had different genomic contents and had evolved in different ways.

The pan-genome's functional annotation uncovered numerous pathways involved in metabolism, transport, and regulation. The presence of genes involved in virulence and antibiotic resistance also varied across the species we examined.

Insights into the evolutionary links among the creatures were gained by building a phylogenetic strain tree based on the pan-genome. Within the strain tree, our data showed separate groups, or clades, which we interpret to represent unique patterns of genomic diversity and evolutionary history. With a few notable exceptions, we found that the evolutionary relationships among the creatures matched their taxonomic classification. The gene count rank distribution ranged from 35 genes with a length of ~6,000 bp. The strain count was observed to be 80. The gene count rank distribution and strain count rank distribution are shown in Figure 1, and the strain tree of the relevant genomes is shown in Figure 2.

## Genomic plasticity analysis

Our research showed that the genomes of the creatures we analyzed contained a wide variety of genomic islands and mobile genetic components. Pathogenicity islands, prophages, and integrative conjugative elements are only a few examples of the genomic islands we found. We also found several different kinds of transposons and insertion sequences, which are mobile genomic elements. Genomic islands and MGEs were annotated functionally, revealing a vast variety of genes involved in virulence, antibiotic resistance, and other cellular activities. Differences in the genomic island and MGE presence across studied taxa suggest distinct genomic content and evolutionary histories. The genome island prediction map is shown in Figure 3.

Our research also found evidence of genetic flexibility in locations that are specific to individual or clade-specific species. Additional research into these regions may be worthwhile because they may house genes with crucial roles in the development of certain phenotypes or functions. Table 4 shows the predicted gene names along with their products.

TABLE 3 Average nucleotide identity of the *Lactiplantibacillus plantarum* genomes under study.

Query genome	DDH	Model C.I.	Distance	Prob. DDH $\geq$ 70%	DDH
GCF_003020005.1_ASM302000v1	82.6	(78.8–85.9%)	0.1192	93.34	91.1
GCF_003020005.1_ASM302000v1	79	(75–82.5%)	0.1373	89.98	88.1
GCF_003020005.1_ASM302000v1	81.4	(77.5–84.8%)	0.1251	92.38	91.6
GCF_003020005.1_ASM302000v1	84.2	(80.5–87.4%)	0.111	94.48	91.5
GCF_003020005.1_ASM302000v1	84.8	(81.1–87.9%)	0.108	94.86	92.2
GCF_003020005.1_ASM302000v1	86	(82.4–89%)	0.1018	95.55	92.3
GCF_003020005.1_ASM302000v1	81.3	(77.4–84.6%)	0.1259	92.24	91.9
GCF_003020005.1_ASM302000v1	81.3	(77.4–84.7%)	0.1256	92.3	92.3
GCF_003020005.1_ASM302000v1	81.8	(77.9–85.2%)	0.1231	92.71	92.6

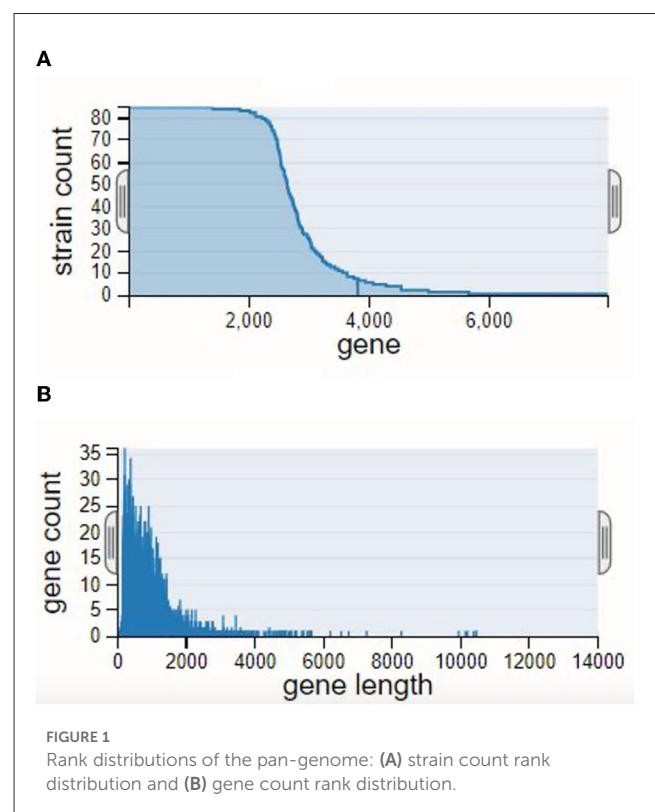
## Prediction of prophage regions, probiotic potential, and bacteriocins

Based on our research, we determined that the genome of the *L. plantarum* K25 strain has four prophage regions. Different prophage families, such as the Siphoviridae and the Myoviridae, were found dispersed throughout the genome. As a further indicator of changes in genomic content and evolution, we found that the prophage areas varied among the taxa we examined. The prophage regions starting from the genes implicated in virulence, antibiotic resistance, and other cellular functions were among 2111511 to 2153545, containing a total of 24 CDS with a GC content of 43.38%. Several functional categories were discovered by functional annotation of the prophage areas. We found that several prophage areas had genes involved in stress response and DNA repair, which raises the possibility that these prophages help protect the organism from environmental challenges. Figure 4 shows the genomic map of the prophage regions present in the genome.

Moreover, our investigation into the genome of the *L. plantarum* K25 strain uncovered prophage remnants, suggesting that the organism may have experienced prophage loss or decay. Insights concerning the organism's evolution and its interactions with other bacteria might be gained by studying these prophage remains. Our research showed that the genome of the *L. plantarum* K25 strain contains numerous clusters of bacteriocin genes. Class I and class II bacteriocins were among the many kinds we found. Additionally, we found that bacteriocin gene clusters varied between the organisms we analyzed, suggesting a spectrum of genomic diversity and evolutionary history.

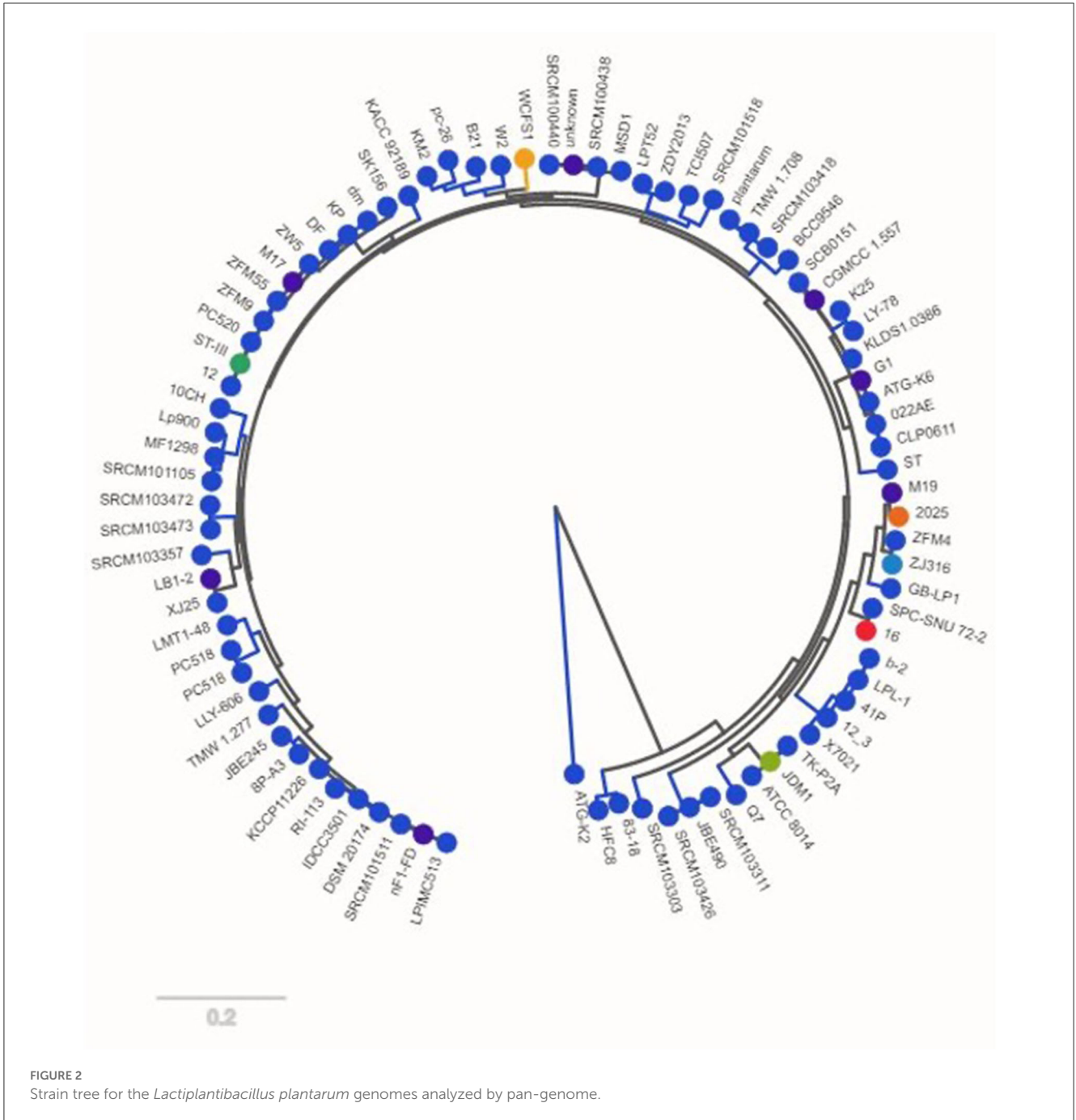
Genes involved in biosynthesis, regulation, and immunity were only some of the many functional categories discovered by the functional annotation of the bacteriocin gene clusters. We found that certain clusters of bacteriocin genes including Enterocin and Plantaricin and antibiotic-resistance genes raise the possibility that bacteriocins function to defend the organism against both environmental challenges and rival germs. Figure 5 shows the map of the bacteriocin genes predicted.

Our research also found several structural variations of bacteriocins in the genome of the *L. plantarum* K25 strain. It may be worthwhile to investigate whether or not these variations have unique antibacterial activities and characteristics.



## Discussion

The present investigation involved a comparative genomics analysis of the K25 strain of *L. plantarum*, aimed at revealing its genomic potential. The results of our analysis indicate that the genome was both compact and efficient, measuring 3.18 Mb in size and possessing a GC content of 43.38%. A comprehensive set of 3,199 protein-coding genes was identified, with 3,120 of them functionally annotated. The results of our analysis indicate a significant level of synteny and conservation of gene order between various strains of *L. plantarum*, suggesting a closely related evolutionary history. A diverse range of genetic components, such as prophage regions and bacteriocin gene clusters, was detected,



indicating prospective prospects in the fields of biotechnology and medicine (An et al., 2023; Davray et al., 2023).

Prior research endeavors on *Lactiplantibacillus plantarum* have predominantly centered on diverse facets of the strain's genome, encompassing the discernment of probiotic characteristics, operational attributes, and genomic heterogeneity (Huang et al., 2023). Liu et al. (2022) conducted a study to examine the genomic and metabolic characteristics of *L. plantarum* JDM1, a strain that was obtained from fermented milk. The research identified various genetic components that play a role in carbohydrate metabolism, such as genes associated with the utilization of lactose and the metabolism of fructose (Echegaray et al., 2022; Liu et al., 2022).

Ge et al. (2020) conducted a study that centered on the genomic diversity and the evolution of *L. plantarum* strains derived from various ecological niches, such as plant-associated and dairy-associated strains. The research successfully identified genetic variations among the strains and proposed that *L. plantarum* exhibits a significant level of ecological adaptation and specialization toward specific habitats (Ge et al., 2020).

In contrast to prior research endeavors, a comparative genomics methodology was employed to detect genetic components that possess prospective utilities in the fields of biotechnology and medicine. These components encompass prophage regions and bacteriocin gene clusters (Kim et al.,

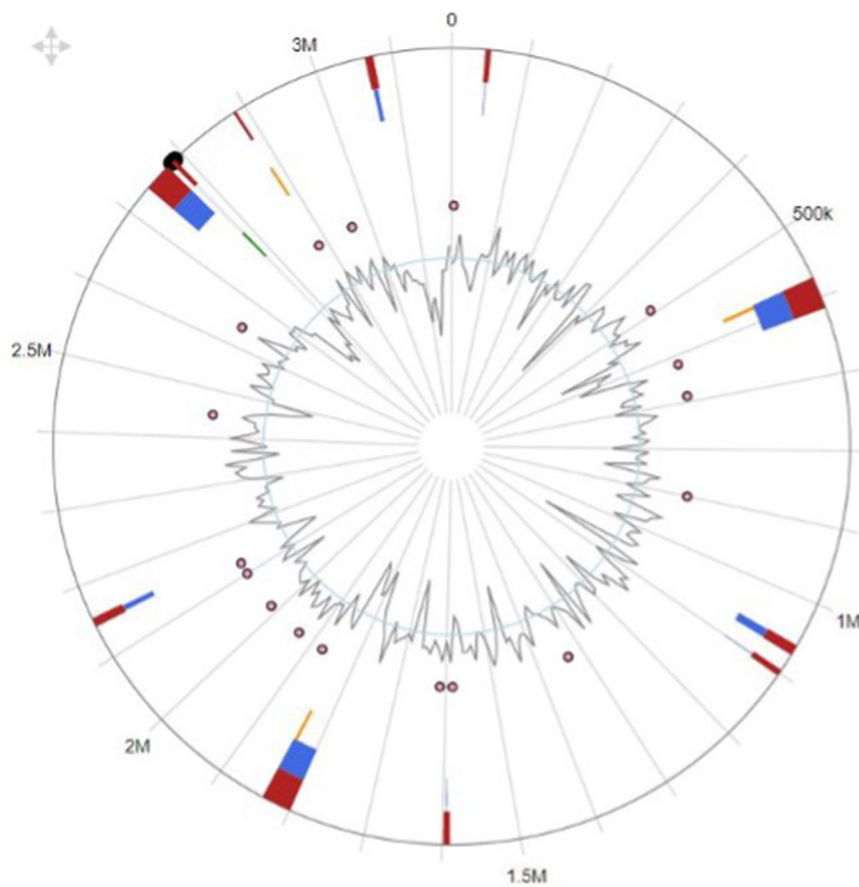


FIGURE 3  
Genomic island overview of the genome plasticity analysis.

TABLE 4 Gene names, accounts, and products computed by genome plasticity analysis.

Gene name	Accnum	Product
B5G53_RS13540	WP_103127150.1	Site-specific integrase
B5G53_RS13545	WP_106904823.1	Hypothetical protein
B5G53_RS13550	WP_106904824.1	Hypothetical protein
B5G53_RS13555	WP_106904825.1	Hypothetical protein
B5G53_RS13560	WP_106904826.1	Hypothetical protein
B5G53_RS13565	WP_106904827.1	Hypothetical protein

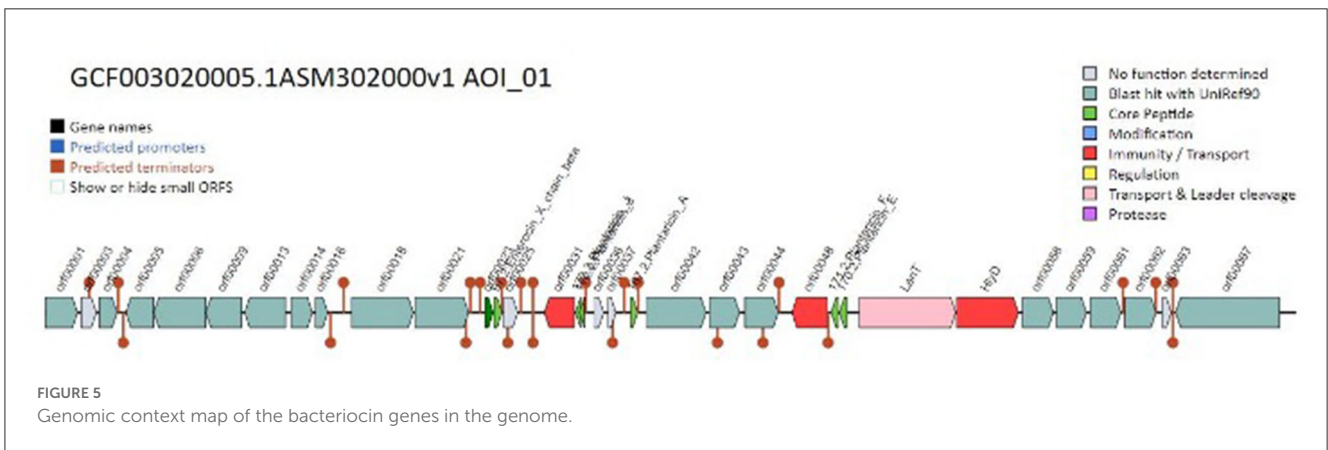
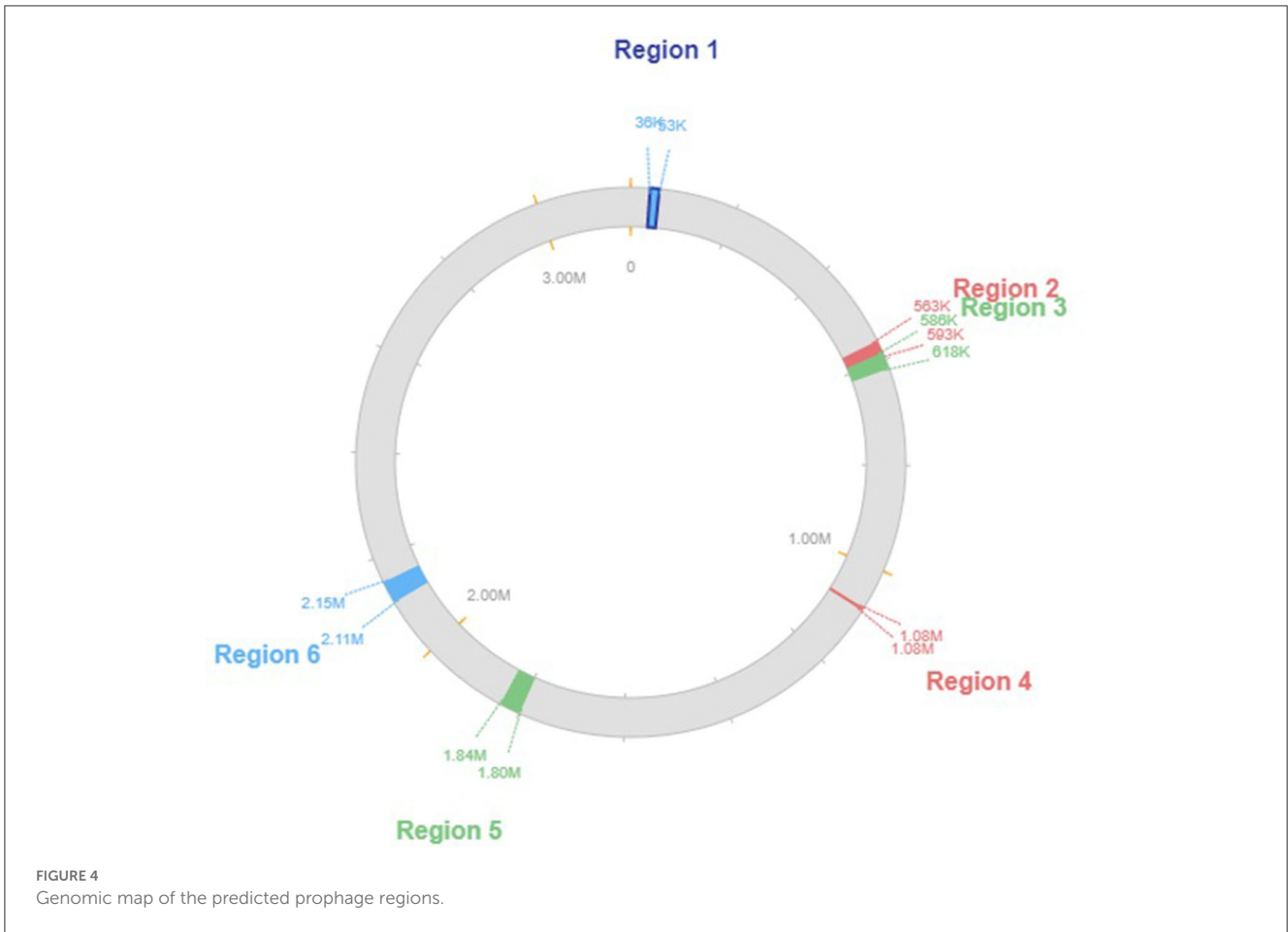
2023). The present study examined the evolutionary connections and genetic preservation of the *Lactobacillus plantarum* K25 strain in comparison to other strains of *L. plantarum*. This analysis offers valuable insights into the evolutionary and adaptive mechanisms of the species in response to diverse ecological habitats (Chokesajjawatee et al., 2020; Aziz et al., 2023).

This study conducted a comparative genomics analysis of the *Lactiplantibacillus plantarum* K25 strain, revealing a

compact and efficient genome with potential applications in biotechnology and medicine. With 3,199 protein-coding genes identified, the analysis demonstrated significant synteny and conservation of gene order among *L. plantarum* strains, suggesting a closely related evolutionary history. The detection of diverse genetic components, such as prophage regions and bacteriocin gene clusters, highlights the potential utility of the K25 strain across various disciplines. This investigation builds upon previous research on *L. plantarum* by providing insights into the genomic capabilities and evolutionary relationships of the K25 strain, offering valuable information for future studies and applications.

## Conclusion

In conclusion, the present investigation has unveiled the fascinating genomic potential of the *L. plantarum* K25 strain through a comparative genomics analysis. Our findings showcase a compact, efficient genome with a rich array of protein-coding genes that underscore the strain's potential applications in biotechnology and medicine. By revealing the remarkable evolutionary connections and genetic conservation among various *L. plantarum* strains, this study not only expands



our understanding of this versatile species but also paves the way for harnessing its potential in diverse fields, ultimately enriching the scientific community’s repertoire of knowledge and innovation. Therefore, the *L. plantarum* K25 strain and the wider field of *L. plantarum* genomes have a very bright future. The findings of this study open new avenues for research, development, and improvement in the fields of biotechnology, medicine, and our general comprehension of this adaptable species. We can expand the boundaries of knowledge, improve human welfare, and progress in several

sectors and fields by making use of the genetic potential of *L. plantarum* K25.

### Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.



## Author contributions

TA, MN, MS, AS, and YZ: conceptualization. TA, MN, AA, MS, and YZ: methodology. MA: software. AA: validation. TA: formal analysis and data curation. TA, MN, MS, and YZ: investigation. YZ, MA, and AA: resources. TA and MN: writing—original draft preparation and writing—review and editing. AA and MS: visualization. YZ: supervision and funding acquisition. AA and MA: project administration. All authors contributed to the article and approved the submitted version.

## Funding

This research study was financially supported by the National Natural Science Foundation of China (Grant No. 31871823) and the National Key Research and Development Program of China (2017YFE0131800).

## References

- An, J., Zhang, Y., Zhao, Z., Huan, R., Yi, H., Wang, H., et al. (2023). Molecular organization and functional analysis of a novel plasmid-borne cps gene cluster from *Lactiplantibacillus plantarum* YC41. *Microbiol. Spectrum* 11, e04150–e04122. doi: 10.1128/spectrum.04150-22
- Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21. doi: 10.1093/nar/gkw387
- Auch, A. F., Klenk, H.-P., and Göker, M. (2010). Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. *Stand. Genomic Sci.* 2, 142–148. doi: 10.4056/sigs.541628
- Aziz, T., Naveed, M., Makhdoom, S. I., Ali, U., Mughal, M. S., Sarwar, A., et al. (2023). Genome investigation and functional annotation of *Lactiplantibacillus plantarum* YW11 revealing streptin and ruminococcin-A as potent nutritive bacteriocins against gut symbiotic pathogens. *Molecules* 28, 491. doi: 10.3390/molecules28020491
- Aziz, T., Naveed, M., Sarwar, A., Makhdoom, S. I., Mughal, M. S., Ali, U., et al. (2022a). Functional annotation of *Lactiplantibacillus plantarum* 13-3 as a potential starter probiotic involved in the food safety of fermented products. *Molecules* 27, 5399. doi: 10.3390/molecules27175399
- Aziz, T., Sarwar, A., Fahim, M., Al Dalali, S., Din, Z. U., Din, J. U., et al. (2020b). In silico characterization of linoleic acid biotransformation to rumenic acid in food derived *Lactobacillus plantarum* YW11. *Acta Biochim. Pol.* 67, 99–109. doi: 10.18388/abp.2020\_5095
- Aziz, T., Sarwar, A., Fahim, M., Al-Dalali, S., Din, Z. U., Din, J. U., et al. (2020a). Conversion of linoleic acid to different fatty acid metabolites by *Lactobacillus plantarum* 13-3 and in silico characterization of the prominent reactions. *J. Chilean Chem. Soc.* 65, 4879–4884. doi: 10.4067/s0717-970220000204879
- Aziz, T., Sarwar, A., Fahim, M., Din, J. U., Al Dalali, S., Din, Z. U., et al. (2020c). Dose-dependent production of linoleic acid analogues in food derived *Lactobacillus plantarum* K25 and in silico characterization of relevant reactions. *Acta Biochim. Pol.* 67, 123–129. doi: 10.18388/abp.2020\_5167
- Aziz, T., Sarwar, A., Naveed, M., Shahzad, M., Shabbir, M. A., Dabool, A. S., et al. (2022b). Bio-molecular analysis of selected food derived *Lactiplantibacillus* strains for CLA production reveals possibly a complex mechanism. *Food Res. Int.* 154, 111031. doi: 10.1016/j.foodres.2022.111031
- Aziz, T., Sarwar, A., ud Din, J., Al Dalali, S., Khan, A. A., Din, Z. U., et al. (2021). Biotransformation of linoleic acid into different metabolites by food derived *Lactobacillus plantarum* 12-3 and in silico characterization of relevant reactions. *Food Res. Int.* 147, 110470. doi: 10.1016/j.foodres.2021.110470
- Barbosa, J., Albano, H., Silva, B., Almeida, M. H., Nogueira, T., and Teixeira, P. (2021). Characterization of a *Lactiplantibacillus plantarum* R23 isolated from Arugula

## Acknowledgments

The authors greatly acknowledge and express their gratitude to the Researchers Supporting Project (Number: RSP2023R335), King Saud University, Riyadh, Saudi Arabia.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- by whole-genome sequencing and its bacteriocin production ability. *Int. J. Environ. Res. Public Health* 18, 5515. doi: 10.3390/ijerph18115515
- Chokesajjawatee, N., Santiyanont, P., Chantarasakha, K., Kocharin, K., Thammamongtham, C., Lertampaiporn, S., et al. (2020). Safety assessment of a nham starter culture *Lactobacillus plantarum* BCC9546 via whole-genome analysis. *Sci. Rep.* 10, 10241. doi: 10.1038/s41598-020-66857-2
- Churro, C., Pereira, P., Vasconcelos, V., and Valério, E. (2012). Species-specific real-time PCR cell number quantification of the bloom-forming cyanobacterium *Planktothrix agardhii*. *Arch. Microbiol.* 194, 749–757. doi: 10.1007/s00203-012-0809-y
- Davray, D., Bawane, H., and Kulkarni, R. (2023). Non-redundant nature of *Lactiplantibacillus plantarum* plasmidome revealed by comparative genomic analysis of 105 strains. *Food Microbiol.* 109, 104153. doi: 10.1016/j.fm.2022.104153
- Dhillon, B. K., Chiu, T. A., Laird, M. R., Langille, M. G., and Brinkman, F. S. (2013). IslandViewer update: improved genomic island discovery and visualization. *Nucleic Acids Res.* 41, W129–W132. doi: 10.1093/nar/gkt394
- Ding, W., Baumdicker, F., and Neher, R. A. (2018). panX: pan-genome analysis and exploration. *Nucleic Acids Res.* 46, e5. doi: 10.1093/nar/gkx977
- Echegaray, N., Yilmaz, B., Sharma, H., Kumar, M., Pateiro, M., Ozogul, F., et al. (2022). A novel approach to *Lactiplantibacillus plantarum*: from probiotic properties to the omics insights. *Microbiol. Res.* 268, 127289. doi: 10.1016/j.micres.2022.127289
- Garcia-Gonzalez, N., Battista, N., Prete, R., and Corsetti, A. (2021). Health-promoting role of *Lactiplantibacillus plantarum* isolated from fermented foods. *Microorganisms* 9, 349. doi: 10.3390/microorganisms9020349
- Ge, Y.-Y., Zhang, J.-R., Corke, H., and Gan, R.-Y. (2020). Screening and spontaneous mutation of pickle-derived *Lactobacillus plantarum* with overproduction of riboflavin, related mechanism, and food application. *Foods* 9, 88. doi: 10.3390/foods9010088
- Gemayel, K., Lomsadze, A., and Borodovsky, M. (2022). MetaGeneMark-2: improved gene prediction in metagenomes. *bioRxiv*. doi: 10.1101/2022.07.25.500264
- Huang, P., Yi, S., Yu, L., Tian, F., Zhao, J., Zhang, H., et al. (2023). Integrative analysis of the metabolome and transcriptome reveals the influence of *Lactobacillus plantarum* CCFM8610 on germ-free mice. *Food Funct.* doi: 10.1039/D2FO03117E
- Jeong, C.-H., Sohn, H., Hwang, H., Lee, H.-J., Kim, T.-W., Kim, D.-S., et al. (2021). Comparison of the probiotic potential between *Lactiplantibacillus plantarum* isolated from kimchi and standard probiotic strains isolated from different sources. *Foods* 10, 2125. doi: 10.3390/foods10092125
- Kim, B., Xu, X., Lee, H., Pathiraja, D., Kim, K., Choi, I.-G., et al. (2023). Complete genome and two plasmids sequences of *Lactiplantibacillus plantarum* L55 for probiotic potentials. *J. Anim. Sci. Technol.* doi: 10.5187/jast.2023.e37

- Liu, G., Liu, Y., Ro, K.-S., Du, L., Tang, Y.-J., Zhao, L., et al. (2022). Genomic characteristics of a novel strain *Lactiplantibacillus plantarum* X7021 isolated from the brine of stinky tofu for the application in food fermentation. *LWT* 156, 113054. doi: 10.1016/j.lwt.2021.113054
- Meier-Kolthoff, J. P., and Göker, M. (2019). TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* 10, 2182. doi: 10.1038/s41467-019-10210-3
- Min, Z., Xiaona, H., Aziz, T., Jian, Z., and Zhennai, Y. (2020). Exopolysaccharides from *Lactobacillus plantarum* YW11 improve immune response and ameliorate inflammatory bowel disease symptoms. *Acta Biochim. Pol.* 67, 485–493. doi: 10.18388/abp.2020\_5171
- Sarwar, A., Aziz, T., Din, J., Khalid, A., Rahman, T., and Daudzai, Z. (2018). Pros of lactic acid bacteria in microbiology: A review. *Biomed. Lett.* 4, 59–66
- Tegopoulos, K., Stergiou, O. S., Kiouisi, D. E., Tsifintaris, M., Koletsou, E., Papageorgiou, A. C., et al. (2021). Genomic and phylogenetic analysis of *Lactiplantibacillus plantarum* L125, and evaluation of its anti-proliferative and cytotoxic activity in cancer cells. *Biomedicines* 9, 1718. doi: 10.3390/biomedicines9111718
- van Heel, A. J., de Jong, A., Song, C., Viel, J. H., Kok, J., and Kuipers, O. P. (2018). BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res.* 46, W278–W281. doi: 10.1093/nar/gky383
- Wang, J., Zhao, X., Tian, Z., Yang, Y., and Yang, Z. (2015). Characterization of an exopolysaccharide produced by *Lactobacillus plantarum* YW11 isolated from Tibet Kefir. *Carbohydr. Polym.* 10, 16–25. doi: 10.1016/j.carbpol.2015.03.003
- Yunyun, J., Y., Zhang, J., Zhao, X., Zhao, W., Yu, Z., et al. (2018). Complete genome sequencing of exopolysaccharide-producing *Lactobacillus plantarum* K25 provides genetic evidence for the probiotic functionality and cold endurance capacity of the strain. *Biosci. Biotechnol. Biochem.* 82, 1225–1233. doi: 10.1080/09168451.2018.1453293
- Zhang, J., Cai, D., Yang, M., Hao, Y., Zhu, Y., Chen, Z., et al. (2020). Screening of folate-producing lactic acid bacteria and modulatory effects of folate-biofortified yogurt on gut dysbacteriosis of folate-deficient rats. *Food Funct.* 11, 6308–6318. doi: 10.1039/D0FO00480D
- Zhang, J., Zhao, W., Guo, X., Guo, T., Zheng, Y., Wang, Y., et al. (2017). Survival and effect of exopolysaccharide-producing *Lactobacillus plantarum* YW11 on the physicochemical properties of ice cream. *Polish J. Food Nutr. Sci.* 67, 191–200. doi: 10.1515/pjfn-2017-0002
- Zhao, D., Wang, Q., Meng, F., Lu, F., Bie, X., Lu, Z., et al. (2022). TetR-type regulator Lp\_2642 positively regulates plantaricin EF production based on genome-wide transcriptome sequencing of *Lactiplantibacillus plantarum* 163. *J. Agric. Food Chem.* 70, 4362–4372. doi: 10.1021/acs.jafc.2c00206