

## Article

# Decoding the Gene Variants of Two Native Probiotic *Lactiplantibacillus plantarum* Strains through Whole-Genome Resequencing: Insights into Bacterial Adaptability to Stressors and Antimicrobial Strength

Gabriela N. Tenea 

Biofood and Nutraceuticals Research and Development Group, Faculty of Engineering in Agricultural and Environmental Sciences, Technical University of the North, Av. 17 de Julio s-21 Barrio El Olivo, Ibarra 100150, Ecuador; gntenea@utn.edu.ec

**Abstract:** In this study, whole-genome resequencing of two native probiotic *Lactiplantibacillus plantarum* strains—UTNGt21A and UTNGt2—was assessed in order to identify variants and perform annotation of genes involved in bacterial adaptability to different stressors, as well as their antimicrobial strength. A total of 21,906 single-nucleotide polymorphisms (SNPs) were detected in UTNGt21A, while 17,610 were disclosed in the UTNGt2 genome. The comparative genomic analysis revealed a greater number of deletions, transversions, and transitions within the UTNGt21A genome, while a small difference in the number of insertions was detected between the strains. A divergent number of types of variant annotations were detected in both strains, and categorized in terms of low, moderate, and high modifier impact on the protein effectiveness. Although both native strains shared common specific genes involved in the stress response to the gastrointestinal environment, which may qualify as a putative probiotic (bile salt, acid, temperature, osmotic stress), they were different in their antimicrobial gene cluster organization, with UTNGt21A displaying a complex bacteriocin gene arrangement and dissimilar gene variants that might alter their defense mechanisms and overall inhibitory capacity. The genome comparison revealed 34 and 9 genomic islands (GIs) in the UTNGt21A and UTNGt2 genomes, respectively, with the overrepresentation of genes involved in defense mechanisms and carbohydrate utilization. In addition, pan-genome analysis disclosed the presence of various strain-specific genes (shell genes), suggesting a high genome variation between strains. This genome analysis illustrates that the bacteriocin signature and gene variants reflect a niche-inherent pattern. These extensive genomic datasets will guide us to understand the potential benefits of the native strains and their utility in the food or pharmaceutical sectors.



**Citation:** Tenea, G.N. Decoding the Gene Variants of Two Native Probiotic *Lactiplantibacillus plantarum* Strains through Whole-Genome Resequencing: Insights into Bacterial Adaptability to Stressors and Antimicrobial Strength. *Genes* **2022**, *13*, 443. <https://doi.org/10.3390/genes13030443>

Academic Editor: Silvia Turrone

Received: 30 December 2021

Accepted: 26 January 2022

Published: 28 February 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the author. 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/>).

**Keywords:** genomic variation; gene variants; SNPs; *Lactiplantibacillus plantarum*; next-generation sequencing; pan-genome; probiotics

## 1. Introduction

Bioprospecting tropical plants to search for beneficial endophytic microorganisms that produce novel biotechnological molecules remains of interest [1,2]. Microorganisms associated with plants are subjected to constant metabolic and environmental interactions; however, the diversity of the produced molecules is linked to the host chemical composition [3,4].

The genome-scale analysis of beneficial bacteria such as lactic acid bacteria (LAB) represents a fundamental approach to investigating their physiological performance or predicting their putative probiotic capabilities, adaptability to different environments (tolerance to bile salt, acids, temperature, osmotic stress, etc.), and post-/metabiotic features (production of antimicrobials, secondary metabolites, enzymes, exopolysaccharides, etc.). Douillard et al. [5] performed a comparative genome analysis of two commercialized

*Lactobacillus casei* strains isolated from different fermented products, showing a limited number of SNPs in the genome, among other features that implied the high similarity among these strains. Moreover, Botta et al. [6] studied the genomes of three *L. plantarum* strains, indicating that the high flexibility and metabolic versatility of these strains—which can acquire, substitute, or delete genomic regions—are ligated to their distinct environmental niches of origin. In another study, three *Lactococcus lactis* isolates showed few mutations within various genes involved in amino acid production/transport and the mismatch repair mechanism (*mutL* gene), highlighting that the high frequency of mutations in this region might be responsible for the adaptation of strains from plant to dairy environments [7]. Among probiotic *Lactobacillus*, the most documented strain is *Lactobacillus rhamnosus* GG (ATCC 53103) [8], isolated from the intestinal tract of a healthy human. Additionally, *L. plantarum* WCFS1, isolated from human saliva, has been used as a reference strain for many genomic studies [9,10]. Previous research indicates that *L. plantarum* strains isolated from various niches exhibit high genetic variation and phylogenetic patterns [11]. In addition, no direct genetic connection between genomic characteristics and host niches was found in the pan-genome analysis [12]. More recent genomic comparison between several *L. plantarum* strains originating from pickles, fermented sauce, and human feces indicated that the genetic variation of LAB strains is associated with the host niche [13]. However, the capacity of microbes to adapt to different niches depends on the genetic repertoire and the capacity of these strains to counteract externally exerted physicochemical challenges [14]. Therefore, the microbial composition in these environments is irregular, and depends on intrinsic (i.e., physical and nutritional conditions) and extrinsic (i.e., environmental and harvesting conditions) parameters of the plant matrix [15]. Although several species have been identified, the selection of new strains with valuable biotechnological properties remains a topic of interest [16].

Native tropical fruits consumed in Ecuador are likely to be an excellent microenvironment to search for such useful microorganisms [17]. These microorganisms produce antimicrobials that enable their survival in competitive habitats with other microorganisms, protecting the host against other bacterial or fungal pathogens [18]. Native lactobacilli from wild fruits regularly face extreme variation in conditions such as temperature, nutrients, and pH; however, their metabolic capacities—including antimicrobial properties—are more competitive than those of other microorganisms isolated from other niches [2]. The biochemical versatility and diversity of these microorganisms suggest that many actively produced molecules remain unknown. Therefore, these probiotics can be an alternative to conventional antibiotics or other therapeutic drugs.

Previously, we prospected the microbiota of several wild fruits of the Ecuadorian Amazon [19] to select potential probiotic and antimicrobial LAB producers. Among these unique biological niches, two strains of *Lactiplantibacillus plantarum*—UTNGt21A and UTNGt2—were selected, and their genomes were characterized. Although they belong to the same species and share conserved genes responsible for the biosynthesis or degradation of structural compounds such as proteins, lipids, and DNA, they are highly divergent, as demonstrated by the differences in their antimicrobial gene cluster organization (i.e., bacteriocins and non-ribosomal peptide biosynthesis gene clusters), as well as their diversity of bioactive molecules and secondary metabolites. In this study, the whole-genome resequencing of both native strains was carried out in order to identify gene variants and discern metabolic features and genes linked to the adaptability of bacteria to different environments. Moreover, the identification of GIs and insertion sequences (ISs) might explain the adaptability, metabolic versatility, and fitness of both native strains. Furthermore, pan-genome analysis was conducted in order to detect strain-specific genes (shell genes). The use of NGS for whole-genome sequencing (WGS) and gene annotation, followed by the identification of the type of SNPs, indels, transitions, and transversions, along with the in-depth evaluation of the antimicrobial cluster gene variants of the target strains and of the reference counterpart (*L. plantarum* WCFS1), might help to understand

their genetic variation, genomic complexity, adaptation to different niches, and overall antimicrobial capacity.

## 2. Materials and Methods

### 2.1. Bacterial Strains

*Lactiplantibacillus plantarum* UTNGt21A and UTNGt2 strains were isolated from wild fruits of *Solanum quitoense* Lam. (naranjilla) and *Theobroma grandiflorum* (white cacao), following the procedure described in [19]. The genome assembly data of the UTNGt2 and UTNUTNGt21A strains were previously deposited to the NCBI database under the BioProject PRIJNA705232 with BioSample SAMN18053630 on 26 February 2021, and BioProject PRIJNA740042 with BioSample SAMN19816459 on June 23, 2021. The assembly of the *Lactobacillus plantarum* WCFS1 strain is available in GenBank ([https://www.ncbi.nlm.nih.gov/assembly/GCF\\_000203855.3](https://www.ncbi.nlm.nih.gov/assembly/GCF_000203855.3), accessed on 5 September 2020), providing the basis for data analysis of the whole-genome sequencing (reference sequence).

### 2.2. Whole-Genome Sequencing, Gene Prediction, and Functional Annotation

The Illumina HiSeq X Ten platform was used for sequencing using a custom assay by the design service (Macrogen Inc., Seoul, Korea). The detailed construction procedure, sequencing library, cluster generation, quality control, and statistical analyses summarized the basic characteristics of the read contig, and assembly was as described in previous studies [19,20]. To map the reads obtained from the sequencing, the Burrows–Wheeler aligner (BWA) (v0.7.17) and Burrows–Wheeler transform (BWT) (<http://bio-bwa.sourceforge.net/>, accessed on 11 September 2020) algorithms were used [21]. The gene prediction and functional annotation were performed as previously described in [19].

### 2.3. Genome and Pan-Genome Comparison Analysis

The circular and linear genome comparison diagram was predicted using CGView [22]. Multiple genome alignment under Mauve (with default settings: the value of minimal Locally Collinear Blocks (LCBs) was equal to 1000; island size: 50; backbone size: 50; maximum gap: 50) was used to perform the synteny analysis [23]. The “Mauve Contig Mover” (default settings) was used to order a draft genome of UTNGt21A and UTNGt2 relative to a related reference genome. Pan-genome analysis was carried out by using Roary v1.007001 [24] with the MAFFT v7.427 aligner [25]. Genomic sequences from the three samples were used to cluster the genes encoding complete protein sequences into core (hardcore and softcore) and accessory (shell and cloud) genomes.

### 2.4. Identification of Genomic Islands (GIs) and Insertion Sequences (ISs) within the Genomes of Native Strains

The webserver IslandViewer 4, used to predict GIs, was employed using WCFS1 as a reference strain [26]. Moreover, to search for ISs, the ISfinder tool (ISsaga V.2.0) was used [27].

### 2.5. SNPs and Indel Discovery, Transition and Transversion Information, and Variant Annotation

Produced mass sequence data were used to search for genetic variation. During mapping, duplicated reads can falsely cause erroneous data to stand out. To prevent this, Sambamba v0.6.7 (<http://lomereiter.github.io/sambamba/>, accessed on 11 September 2020) was used to remove duplicated reads that were identified using mapping information such as start position and CIGAR string (Compact Idiosyncratic Gapped Alignment Report) [28]. After removing duplicates and identifying variants with SAMtools (<http://samtools.sourceforge.net/>, accessed on 11 September 2020) [29,30], the information of each variant was gathered and classified by chromosomes or scaffolds. The numbers of transitions (Ts) and transversions (Tv), along with the Ts/Tv ratio, were calculated using the base change count. Base changes (DNA substitutions) are of two types: interchanges of purines (A <-> G) or pyrimidine (C <-> T) are transitions, while interchanges of a purine for

pyrimidine bases, and vice versa, are transversions. To determine the annotation information, such as amino acid changes of variants, SnpEff v4.3t (<http://snpeff.sourceforge.net/>, accessed on 11 September 2020) was used [31]. SnpEff generates the genes and transcripts affected by the variant, the location of the variants, and how the variant affects the protein synthesis (e.g., generating a stop codon). Because genes usually have multiple transcripts, a single variant can have different effects on different transcripts. Thus, the impact categories must be used with care, as they were created only to help and simplify the filtering process. A representative transcript was chosen by the gene name obtained from the variant calling analysis. Other transcripts were chosen by the information of neighboring genes that were close enough.

### 2.6. Detection of Biosynthetic Gene Clusters (BGCs) and Genes Involved in the Adaptability to Several Stressors

The detection of biosynthetic gene clusters was investigated using the BAGEL4 (<http://bagel.molgenrug.nl/>, accessed on 4 October 2021) webserver [32]. The genes related to the adaptability to pH, bile salt hydrolase, temperature, and osmosis were retrieved from EggNOG annotation results. Moreover, comparison of gene variants was performed using the gene annotations obtained in Section 2.5.

## 3. Results and Discussion

### 3.1. Comparative Genome and Pan-Genome Analysis Reveals the High Genetic and Niche-Specific Variation of Native Strains

To map the reads obtained from sequencing, *Lactobacillus plantarum* WCFS1 was used as a reference genome. For UTNGt21A, 13,121,820 total reads were mapped with 43.12 GC% and Q30 of 94.25%, while for UTNGt2, 11,733,026 total reads were mapped with 43.78 GC% and Q30 of 93.99%. The number of mapped sites, mapping coverage, total number of reads, number of mapped reads, overall mapping ratio, number of mapped bases, and the average alignment depth are shown in Table 1. A circular and linear map comparison is depicted in Figure S1, the UTNGt21A genome being larger than the reference WCFS1 and UTNGt2 genomes. Whole-genome alignments performed using the Mauve contig mover module indicated more rearrangements for the UTNGt21A strain (Figure 1A) than for the UTNGt2 (Figure 1B) strain when both genomes were compared with the reference WCFS1 genome. Although overall high synteny conservation was observed among the three strains, indicating that *L. plantarum* has a very stable genomic structure, numerous gene rearrangements were detected, with highly variable regions located between 1000 kb and 2500 kb (Figure 1A,B). The intersection and connection lines indicate the presence of some regions without LCB outlines, suggesting the presence of strain-specific regions; these differences might be related to the differences in the lifestyle islands (Table S1). These events can occur during evolution or horizontal gene transfer (HGT), indicating the plasticity of the *L. plantarum* genome [33]. Moreover, a total of 4074 genes were subjected to the pan-genome analysis (Figure 2A). The results indicated that the three genomes shared ~2498 common genes, while 1576 genes were sample-specific genes (shell genes). Several gene clusters were distinguished in both native strains (Figure 2B). From the gene presence/absence comparison, several strain-specific proteins involved in the defense mechanism were detected. Among them, a nisin leader peptide-processing serine protease NisP (*nisP*, locus tag: UTNGt21A\_02824) indicated 100% sequence identity with peptidase S8 specific protease domain—a lantibiotic (lanthionine-containing antibiotics)-specific protease very similar in structure to serine proteases from several *L. plantarum* strains and *Bacillus* spp. Likewise, two hypothetical proteins of the lanthionine synthetase C family (COG4403) were detected, with 100% identity with the DUF4135-domain-containing protein found in bacterial species and archaea, of ~380 amino acids in length and unknown function. Moreover, a hypothetical plantaricin-C-like protein (locus tag: UTNGt21A\_02827) showed 100% identity with a plantaricin-C-like precursor from different *L. plantarum* and an autolytic lysozyme (*Lys\_2*) that was annotated in the UTNGt21A genome. BLASTN protein analysis indicated a specific hit with the bacterial SH3 domain and a non-specific

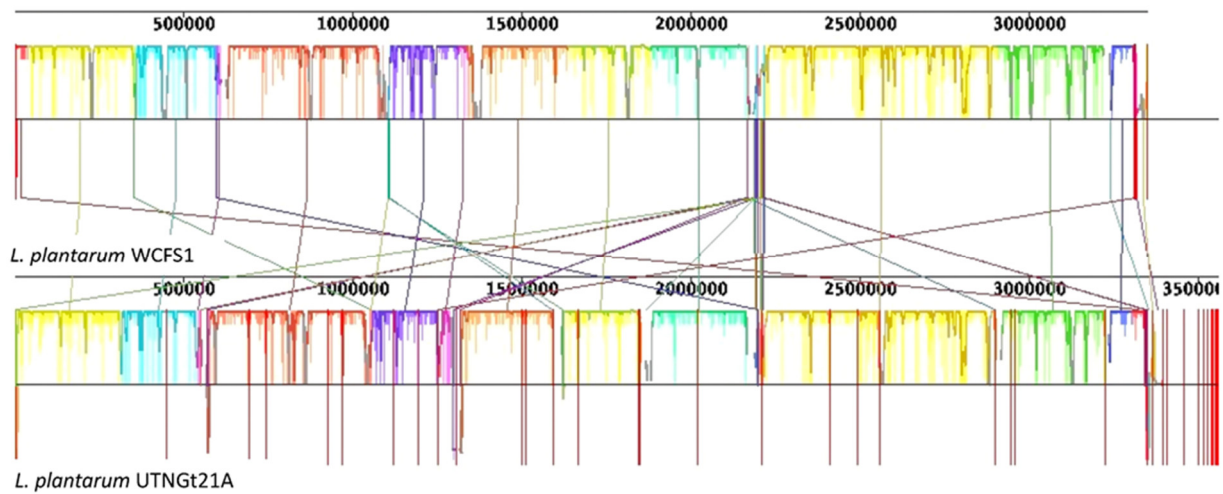
hit with GH25\_Lys2-like—a cell wall endolysin produced by *L. fermentum*. This peptide degrades bacterial cell walls by catalyzing the hydrolysis of 1,4-beta-linkages between N-acetylmuramic acid and N-acetyl-D-glucosamine residues [34]. Likewise, CRISPR-associated endonucleases Cas1, Cas2, and Cas9 were annotated in the UTNGt2 genome, but not in the reference WCFS1 and UTNGt21A genomes. BLASTN protein analysis of CRISPR-associated endonucleases Cas2 and Cas1 indicated 100% identity with the type-II-A-CRISPR protein Csn2 and CRISPR-associated endonuclease Cas2, respectively, of many *L. plantarum* strains. The presence of these endonucleases might enhance the strains' stability and adaptation to new niches [35]. Two genes encoding for an arsenical resistance operon repressor (*arsR*) and arsenate reductase (*arsC*) were annotated in all target genomes, while the genes *arsD*, *arsA*, and *arsB* encoding for the arsenite efflux transporter metallochaperone ArsD, arsenical pump-driving ATPase, and arsenical efflux pump membrane protein ArsB, respectively, were detected in the WCFS1 genome. These results are consistent with previous gene analysis indicating that WCFS1 harbors a protein system involved in arsenic detoxification [36]. Moreover, genes encoding for the biosynthesis of riboflavin synthase (*ribE*) and riboflavin biosynthesis protein RibBA (*ribBA*) were detected in all genomes, while riboflavin biosynthesis protein RibD (*ribD*) and riboflavin transporter RipZ (*ripZ\_2*) were annotated in the genomes of both native strains. The gene encoding for riboflavin transporter RibU (*ribU*) was annotated in the UTNGt21A genome. Early studies indicated that several *Lactobacillus* strains harbor the operon for riboflavin biosynthesis [37]. Moreover, we found that the gene encoding for glutamate decarboxylase (*gadB*) was common to all strains, while the gene encoding for aspartate-1 decarboxylase (*panD*) was detected in the UTNGt21A and WCFS1 genomes, but not UTNGt2. Although from gene analysis we cannot predict the impact of these genes on the production of biogenic amines (toxins that, when accumulated in food during storage, might cause human health problems) by the LAB strains, previous in vitro analysis indicated that these decarboxylase enzymes are not sufficient to produce these substances [38]. However, the pan-genome results indicated a high genomic variation among the native strains, which might be related to the addition or deletion of the genes during the adaptation of the species to the fruit microenvironment. Further transcriptomic and proteomic analysis will elucidate the physiological importance of these genes in bacterial survival, pathogen exclusion, and adaptation to different environments.

**Table 1.** Mapping data stats.

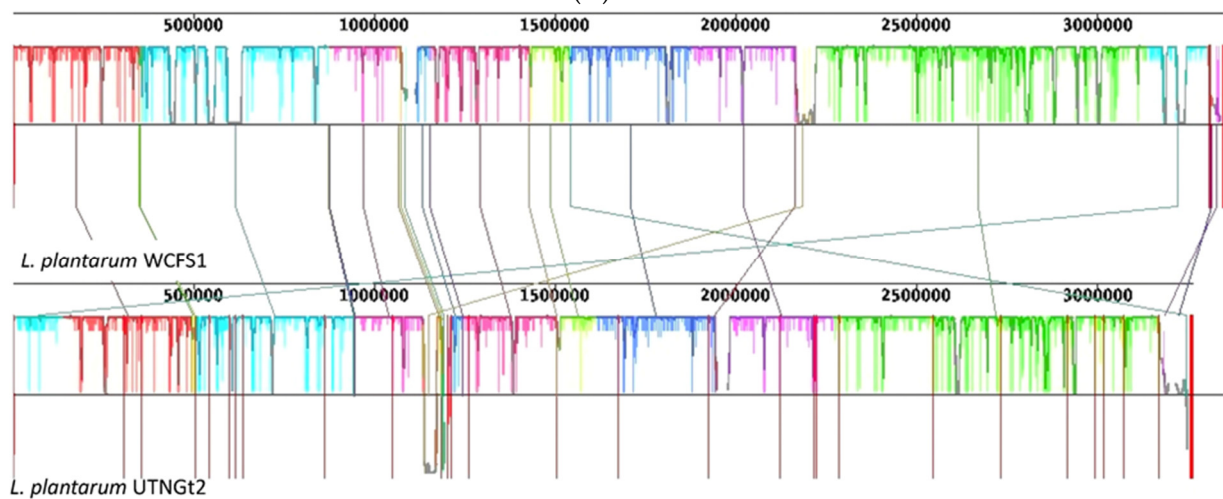
Library Name	Ref. Length	Mapped Sites (≥1x)	Total Reads	Mapped Reads	Mapped Bases	Mean Depth
UTNGt21A	3,348,624	3,028,007 (90.43%)	13,121,820	10,291,789 (78.43%)	941,582,288	281.18
UTNGt2	3,348,624	2,979,050 (88.96%)	11,733,026	10,428,975 (88.89%)	957,588,188	285.96

Note—Library name: sample name; Ref. Length: length of reference genome; Mapped Sites: length of mapped site; Total Reads: number of the total reads; Mapped Reads: number of reads mapped to the reference; Mapped Bases: number of bases in reads mapped to the reference; Mean Depth: average alignment depth.



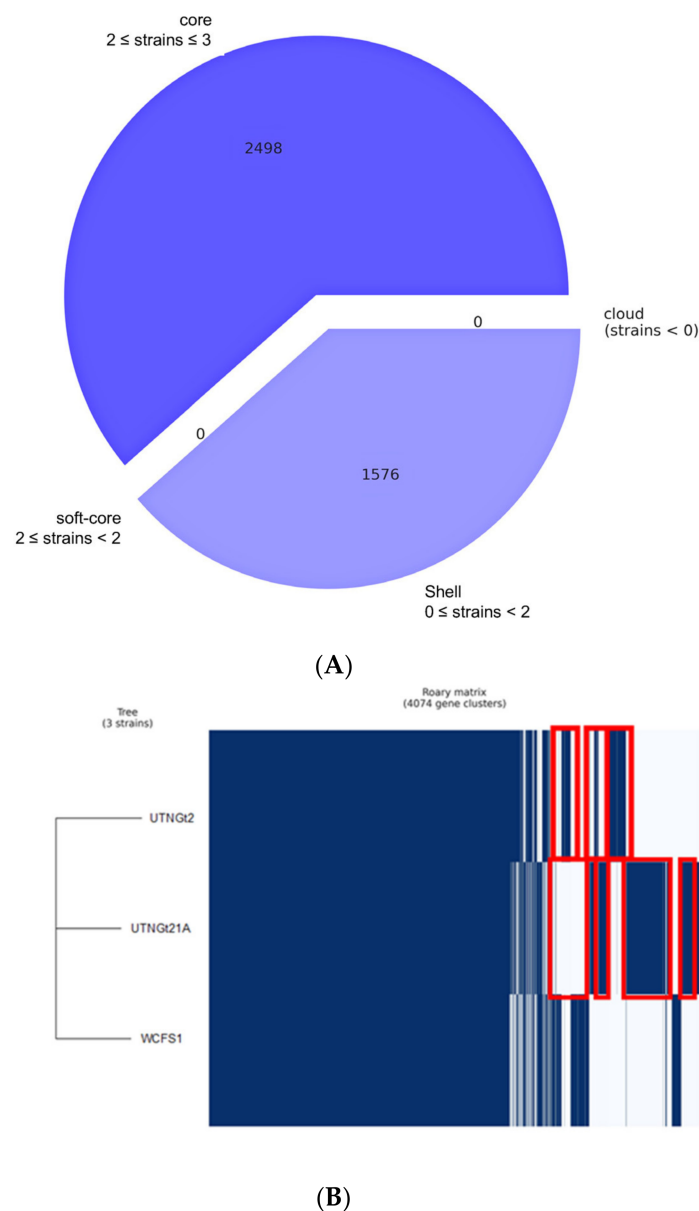


(A)



(B)

**Figure 1.** Whole-genome alignments performed using the Mauve contig mover module. The genome of reference strain WCFS1 was aligned with the (A) UTNGt21A and (B) UTNGt2 draft genomes. White areas indicate low-identity regions between strains. Regions with the same color indicate high-similarity syntenic blocks, and are connected by the same color bars. Red bars indicate boundaries of the original contigs. The numbers above the alignments indicate the nucleotide positions in the WCFS1 genome.

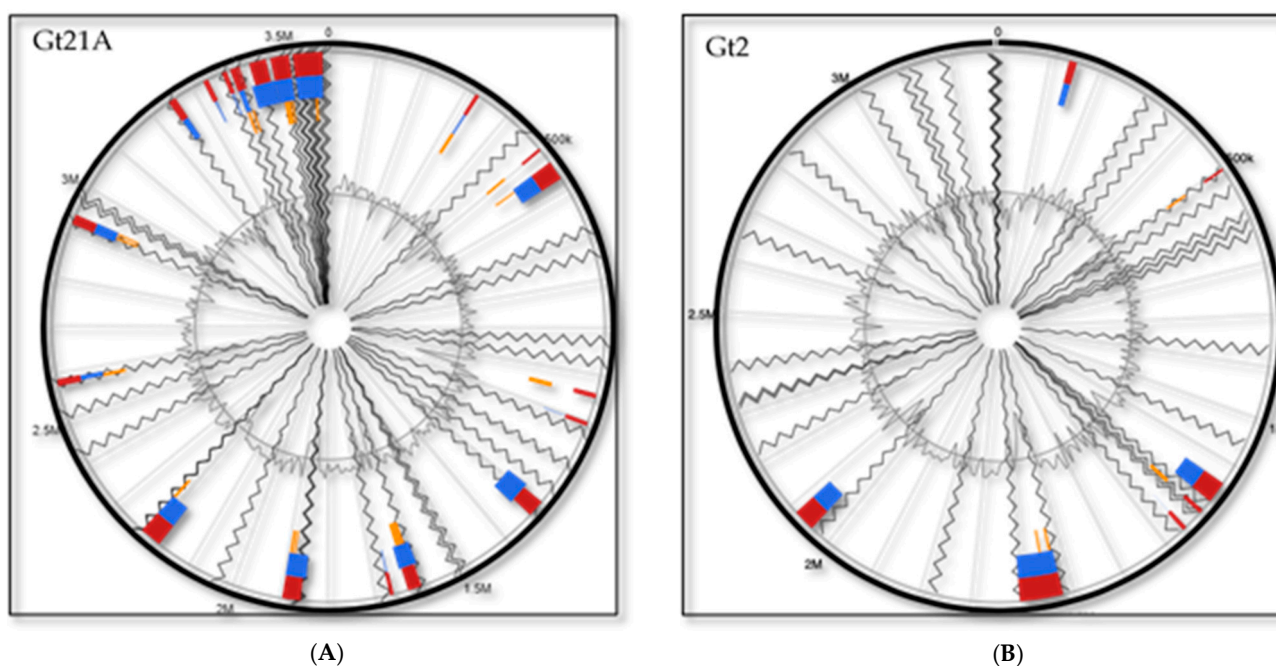


**Figure 2.** Gene content comparison of the three *L. plantarum* strains: (A) The pie chart shows the number of genes belonging to the core, the soft core, the shell, or the cloud. (B) The Roary matrix shows genes typical of each strain and those conserved in all strains. The red rectangular frame indicates several distinct cluster genes of the UTNGt21A and UTNGt2 strains.

### 3.2. Differences between the GIs and ISs Might Explain the Strains' Adaptability to Different Niches

Comparative genomic analysis of native *L. plantarum* strains allows for the identification of niche- or lifestyle-specific genome characteristics. However, acquisition of the host bacterial genomes of bacteriophages, transposons, and other mobile elements via HGT results in the formation of genomic islands (GIs), which might confer fitness benefits on the native strains in specific habits [39]. Genomic analysis via the IslandViewer 4 web tool revealed that the UTNGt21A genome harbored the largest number of islands (34), with a total length of 549074bp, while UTNGt2 harbored 9 GIs with a total length of 214,624bp (Figure 3). In this study, most GIs encoded several hypothetical proteins and species-specific proteins. Thus, within the UTNGt21A genome, several GIs encoding proteins involved in defense mechanisms such as lactococcin-G-processing and transport ATP-binding protein

(LagD) and nisin leader peptide-processing serine protease (NisP) were detected (Table S1). Like the results from the pan-genome analysis, GIs identified sequences matching an autolytic lysozyme (*lys\_2*) and two caseinolytic protease-encoding genes (*ClpP1\_1*, *Clp1\_2*). These bacteriocin-specific genes might enhance the adaptability and competitiveness of the microbe within a niche and, thus, might contribute to the general environmental adaptation of the strain. Within the UTNGt2 genome, GIs encode several proteins involved in the carbohydrate transport PTS system component (e.g., sorbose-specific EIIA, EIIB, EIIC, EIID), PTS system galactitol-specific EIIB component, PTS system N,N'-diacetylchitobiose-specific EIIC component, and fructose-1,6-bisphosphatase class 3. A relatively high number of genes related to carbohydrate utilization, known as “lifestyle adaptation islands”, were previously found in several *L. plantarum* strains [40]; likewise, several hypothetical proteins were found in the GIs of the UTNGt2 strain. In addition, a putative transposon (Tn552 DNA-invertase bin3) and a serine recombinase (PinR) were detected in the GIs of the UTNGt2 genome, while an SPBc2-prophage-derived glycosyltransferase (SunS) was detected within the UTNGt21A genome. No virulence factors or antibiotic-resistance genes were annotated within the GIs. The presence of the PTS carbohydrate system in the UTNGt2 genome might explain the adaptation of the cells to various food matrices. Recent in vitro complementary analysis indicated that the UTNGt2 strain adapted and grew in both milk- and fruit-based matrices, while the UTNGt21A strain adapted and grew in fruit but not dairy matrices (data not shown).



**Figure 3.** Circular plots visualizing genome islands predicted in the (A) UTNGt21A and (B) UTNGt2 genomes, aligned against the complete reference genome *L. plantarum* WCFS1, with blocks colored according to the prediction method. IslandPath-DIMOB (blue), SIGI-HMM (orange), and the integrated results (dark red).

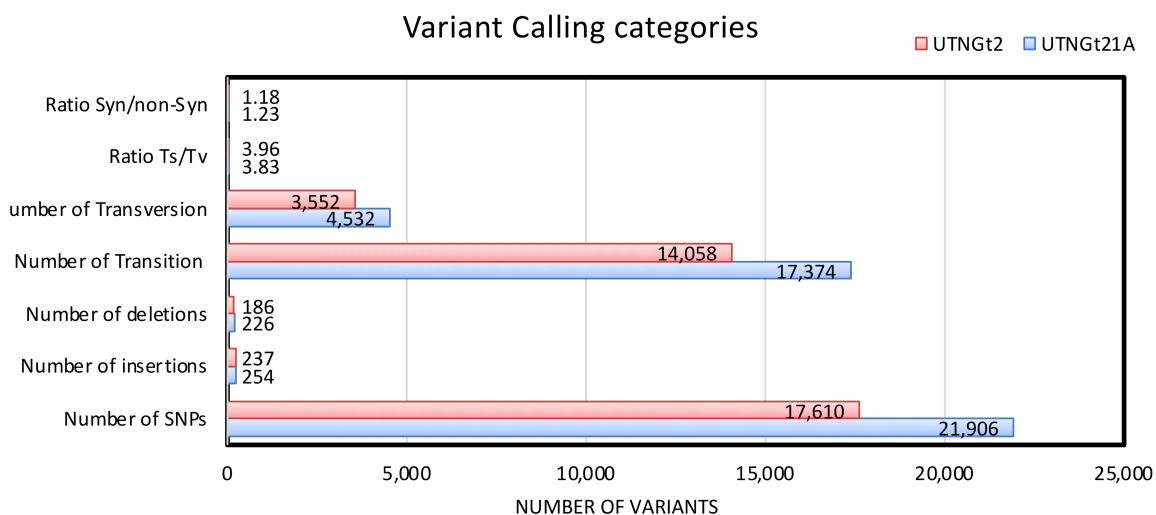
Moreover, several insertion sequences (ISs) were annotated with EggNOG and the ISfinder web tool. A total of 13 and 25 ISs were predicted in different loci on the UTNGt2 and UTNGt21A genomes grouped in eight families. The types of IS and their distribution in both genomes are depicted in Figure S2. Early genome comparison analysis between *L. casei* ATCC334 and other sequenced lactobacilli revealed a relatively high number of IS elements and carbohydrate-related genes [40]. The UTNGt2 genome harbors two distinct transposases—IS256 and ISNCY—while UTNGt21A showed three different transposases, IS3\_ssgr\_IS3, IS1182, and IS5\_ssgr\_IS1031 (identified members of the IS family according



to ISSaga annotation). These genomic datasets are likely to provide answers about strains' adaptation and ecological strength, as well as their role in distinct environments.

### 3.3. Detection of Variants (Insertions, Deletions, SNPs) and Their Impact on the Genomic Architecture

The gene variants (SNPs and short indels) were identified by analyzing the information taken from aligned reads. The variants were classified by each chromosome or scaffold, and the information of the location was marked. Although similar reads were mapped (10,191,789–10,428,975) for both strains, 22,386 and 18,033 variants were detected in the UTNGt21A and UTNGt2 genomes, respectively. Table S2 discloses the base changes for every SNP relative to the reference strain. The summary of the variant calling for each sample is depicted in Figure 4. Moreover, the numbers of transitions (Ts) and transversions (Tv), along with the Ts/Tv ratio, were calculated according to the base change count. Although there are twice as many possible transversions, transitions are more common than transversions due to differences in structural characteristics. The Ts/Tv ratio between homologous strands of DNA is generally ~2.0, but it is typically elevated in coding regions where transversions are more likely to change the underlying amino acid, thus possibly leading to a fatal mutation in the translated protein [41]. For UTNGt21A, the ratio percentage of Ts/Tv was 3.96 and 3.82 for UTNGt2, while the synonymous variants/non-synonymous variants ratio was 1.18 and 1.23 for UTNGt21A and UTNGt2, respectively. Generally, transitions are less likely to result in amino acid substitutions, thus remaining as "silent substitutions" in populations as SNPs, while transversions are more likely to cause amino acid sequence changes. The numbers of SNPs, Ts, and Tv were greater within the UTNGt21A genome. A total of 21,906 SNPs were detected in the UTNGt21A genome, while 17,610 were detected in the UTNGt2 genome. We speculated that the higher number of SNPs might be related to the larger genome of the UTNGt21A strain (3.82 Mpb). Table 2 depicts the number of variants and a brief description of the top 10 types of annotations. The synonymous variant acquires a maximum ratio of 71.51 and 72.01 within the UTNGt21A and UTNGt2 genomes, respectively, which is indicative of a sequence variant where there is no resulting change to the encoded amino acid.



**Figure 4.** Cumulative summary of variant calling categories detected in the UTNGt21A and UTNGt2 genomes. Legend—Library name: sample name; Number of SNPs: number of SNPs in sample; Number of insertions: number of insertions in sample; Number of deletions: number of deletions in sample; Number of Transition (Ts): number of transitions in sample; Number of Transversion (Tv): number of transversions in sample; Syn: number of synonymous variants in sample; non-syn: number of non-synonymous variants in sample.

**Table 2.** Cumulative types of variant annotation, descriptions, and impacts.

Type of Variant Annotation	Description	Impact *	Library Name			
			UTNGt21A		UTNGt2	
			Count	Ratio (%)	Count	Ratio (%)
synonymous_variant	Variant causes a codon that produces the same amino acid (e.g., Ttg/Ctg, L/L)	Low	12,350	71.51	9772	72.01
missense_variant	Variant causes a codon that produces a different amino acid (e.g., Tgg/Cgg, W/R)	Moderate	4740	27.45	3664	27
frameshift_variant	Insertion or deletion causes a frame shift (e.g., an indel's size is not a multiple of 3).	High	61	0.35	33	0.24
stop_gained	Variant causes a STOP codon (e.g., Cag/Tag, Q/*)	High	32	0.19	21	0.16
splice_region and stop_retained_variant	A sequence variant in which a change has occurred within the region of the splice site, either within 1–3 bases of the exon or 3–8 bases of the intron/Variant causes stop codon to be mutated into another stop codon (the new codon produces a different AA). (e.g., Atg/Ctg, M/L (ATG and CTG can be START codons))	Low	18	0.1	22	0.15
conservative_inframe_deletion	One or many codons are deleted (e.g., a deletion multiple of three at a codon boundary).	Moderate	15	0.09	10	0.15
disruptive_inframe_insertion	One or many codons are inserted (e.g., an insertion multiple of three at a codon boundary).	Moderate	12	0.07	9	0.08
stop_lost and splice_region_variant	Variant causes stop codon to be mutated into a non-stop codon (e.g., Tga/Cga, */R)/A sequence variant in which a change has occurred within the region of the splice site, either within 1–3 bases of the exon or 3–8 bases of the intron.	High	9	0.05	5	0.07
disruptive_inframe_deletion	One codon is changed and one or many codons are inserted (e.g., an insert of a multiple of three in size, not at a codon boundary).	Modifier	6	0.03	6	0.04
conservative_inframe_insertion	Inversion of a large chromosome segment (over 1%, or 1,000,000 bases).	Moderate	6	0.03	6	0.04
non_coding_transcript_exon_variant	Region that does not code for any protein or does not carry genetic code.	Low	0	0	6	0.04

Note: \* Impact—High: the variant is assumed to have a high (disruptive) impact on the protein, probably causing protein truncation, loss of function, or triggering nonsense-mediated decay; Moderate: a non-disruptive variant that might change protein effectiveness; Low: assumed to be mostly harmless or unlikely to change protein behavior; Modifier: usually noncoding variants or variants affecting noncoding genes, where predictions are difficult or there is no evidence of impact.

A total of 32 (0.19%) and 21 (0.15%) stop-gained variants of genes involved in the defense mechanism were annotated within the UTNGt21A and UTNGt2 genomes, respectively. A stop\_lost (c.553T>G; p.Ter185Gluxet) and stop\_gained (c.150C>A; p.Cys50) mutation trigger the *tnpR1* locus within both native strains. These genes encode a serine recombinase protein (resolvase) that catalyzes the site-specific recombination of the transposon and regulates its frequency of transposition. A genomic evolutionary study of some *Lactococcus* strains indicated the adaptation of the strains to the milk culture medium because of the loss of the mobile elements [7]. In addition, the UTNGt2 genome harbors six noncoding\_transcript\_exon\_variant counts detected in the *lp\_rRNA* locus (transcribing 16S rRNA). Although the bacteria harbor far fewer noncoding exon variants, their function remains unknown. According to early research, noncoding exons are functionally interchangeable, with alternative splicing generating a larger number of potential regulatory RNAs and an enormous transcriptional repository for gene evolution [42]. Moreover, a stop\_lost mutation was detected within the *ISP2\_1* gene (c.1653A>C; p.Ter551Tyrext) of the UTNGt21A genome. This protein showed a specific hit (E-value:  $8.17 \times 10^{-16}$ ) on COG3666

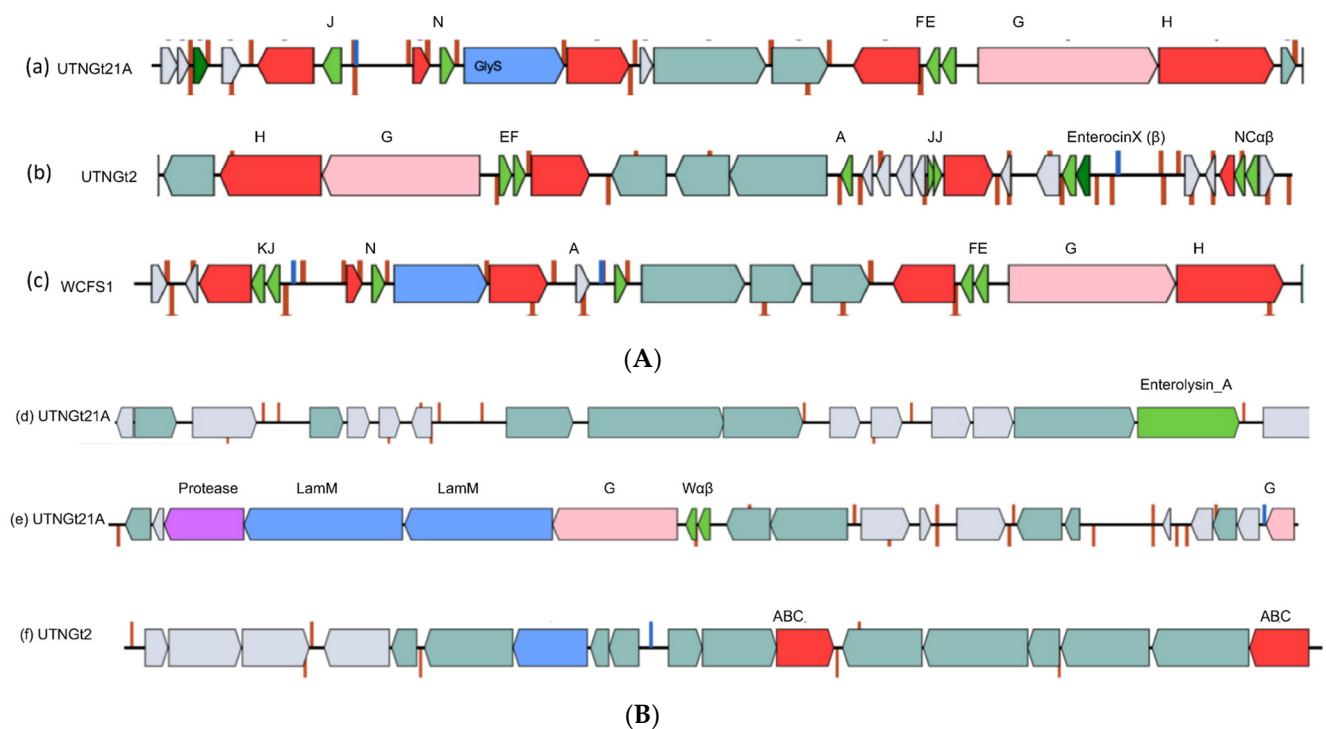
transposase and the transposase DDE domain (pfam01609) (E-value:  $1.58 \times 10^{-16}$ ) protein required for efficient DNA transposition [43]. These transposases were detected in several *L. plantarum* strains, indicating the complexity of the genome and genetic variation of these species [43]. However, these variants might contribute to the overall genomic plasticity and genetic variability, offering occasional mutations to cope with environmental changes and adapt to novel environmental niches.

Additionally, 61 and 33 frameshift\_variants were annotated within the UTNGt21A and UTNGt2 genomes, respectively. Among them, some strain-specific variants were detected within genes involved in carbohydrate metabolism and adherence-associated mucus-binding proteins. For example, a deletion was detected within the *pst11A* (c.13\_16delAAAA; p.Lys5fs), *dprA* (c.134delC; p.His46fs), and *aapA* (c.1865delA; p.Gln622fs) genes in the UTNGt21A genome only. The PTS system is essential for sugar metabolism [3], while DprA protects against incoming foreign DNA [44]. AapA is an adherence-associated mucus-binding protein, which contributes to the persistence of bacteria in the human gut, thus exerting probiotic effects [45]. The frameshift, stop\_gained, stop\_lost, and splice\_region variants cause gene expression changes, and may affect the functional properties of the encoded protein when the mutation is located within the coding region. SnpEff reports putative variant impact to make it easier and faster to categorize and prioritize variants. However, the impact categories must be used with care, as they were created only to help simplify the filtering process. Nevertheless, we cannot predict whether a high, moderate, modifier, or lower impact variant is generating a phenotype of interest; further *in vitro* analyses are required to demonstrate this statement.

### 3.4. BGC Organization and Detection of Gene Variants Might Explain the Inhibitory Strength of the Native Strains

The plantaricin genes are arranged in operons *plnABCD*, *plnEFI*, *plnJKLR*, *plnMNOP*, and *plnGHSTUVWX* [3]. Genome annotation analysis revealed the presence of a common two-peptide bacteriocin (*plnEF*) and several genes encoding for the ABC transport system. The genomic organization of the *pln* locus of the three strains is depicted in Figure 5A. Based on the EggNOG results, the *plnA* gene showing 100% identity with the reference gene (locus lp\_0415) was annotated in the UTNGt2 genome [16], but not the UTNGt21A genome. Nevertheless, 12 *plnA* downstream variants (located at 3' of gene) were annotated in the UTNGt21A genome, while 20 upstream variants were detected in the UTNGt2 genome. The types of variants annotated, and the corresponding products, are shown in Table S3. In addition, a greater number of *plnB*, *plnC*, and *plnD* variants were annotated in the genome of UTNGt2, with the most abundant missense variant (15 counts) of the *plnB* gene. The impact of these mutations on gene product function or inhibitory capacity cannot be assessed by gene analysis. Nevertheless, the antimicrobial capacity of the UTNGt2 strain was demonstrated by different molecular assays [16]. Moreover, a frameshift variant (TGG insertion) was detected within the *plnJ* gene of UTNGt21A (c.154\_155insC; p.Ile53fs), but not UTNGt2 (Table S3). In addition, enterolysin\_A, protease, plantaricin W (alpha and beta), and LanM were detected in the UTNGt21A genome only (Figure 5B). A putative enterolysin\_A was previously detected in the genome of *Weissella cibaria* strain UTNGt21O [17], sharing the same ecological origin as the UTNGt21A strain. Moreover, multiple genome sequence alignment was performed with Jalview (version 2.10.1) [46], and the average distance was calculated based on the percentage of sequence similarity between enterolysin\_A from UTNGt21A and the sequences from UTNGt21O and *Enterococcus faecalis* (NCBI accession no. AGG79281.1) strains (Figure S3A,B). However, the UTNGt21A and UTNGt21O showed 50% sequence identity; thus, we speculated that this bacteriocin could be acquired during a horizontal gene transfer between species inhabiting the same microenvironment, but additional analyses are needed in order to confirm this statement. No gene variants were detected, as the resequencing genome annotation was based on the available genomic information of the reference strain. Two hypothetical proteins of class IIb bacteriocin—a hypothetical protein encoding the plantaricin NC8 alpha-peptide

precursor (97.87% identity to *L. plantarum* TBX52118.1) and the plantaricin NC8 beta-peptide precursor (100% identity to *L. plantarum* subsp. *plantarum* NC8)—were annotated within the UTNGt2 genome only. In addition, within contig 4 of UTNGt2, two ABC bacteriocin transporters were found (Figure 5B). Similarly, within the UTNGt21A genome, several species-specific lactococcin-G-processing and transport ATP-binding proteins were predicted. In a recent complementary inhibitory analysis, we evaluated the effects of the peptide extracts from both native strains on the whole-protein profile of *Staphylococcus aureus* ATCC1026. The results indicated a divergent protein pattern as a result of the effect of the peptide treatment on the release of several low- and high-mass proteins (Figure S4). These results are consistent with the findings of our previous study [47], suggesting that the overall inhibitory capacity depends on the peptide–protein extract mixture released in the extract, and is interconnected with the strains’ bacteriocin-encoding repertoire and gene variants, along with their distinct molecular mechanism of action. Decoding the antimicrobial capability via both in silico and in vitro analysis, as well as further coupling with ex vitro evaluation of the inhibitory action, will help to prospect their use as alternatives to conventional antibiotics.



**Figure 5.** (A) Genetic organization of the *pln* loci of (a) UTNGt21A; (b) UTNGt2; and (c) WCFS1. (B) Additional bacteriocin cluster genes detected on (d, e) UTNGt21A and (f) UTNGt2. Legend—A: *plnA*; J: *plnJ*; N: *plnN*; K: *plnK*; E: *plnE*; F: *plnF*; NCαβ: NCα and NCβ; enterocinX (β): enterocin X (chain beta); GlyS: glycotransferase family 2 protein (PInO), LamM: lantibiotic mersacidin-modifying enzyme; Wαβ: Plantaricin\_W (alpha), Plantaricin W (beta); 183.2; Plantaricin\_W (beta); ABC: ABC-type bacteriocin transporter. Red blocks: immunity and transport; green arrow: core peptide.

### 3.5. Gene Variants Might Play Important Roles in the Strains’ Adaptability to Different Stressors and Overall Probiotic Performance

The adaptation capabilities are species-specific and correlate with the genetic repertoire and gene variants of the target strain. The probiotic features of lactobacilli require the survival of the target strain under several stress conditions—including acidic environment, bile salt, and osmotic conditions—along with good viability that must be maintained during product manufacturing conditions, such as temperature and oxidative stress [48]. Moreover, the strains must meet the safety requirements [49]. However, as any changes

in gene and protein expression might occur during severe conditions, the strains might possess several protectors' molecules [50].

Survival under acidic conditions is one of the critical parameters of LAB strains if intended to be used as probiotic cells. The key regulator of intracellular pH is the F0F1-ATPase system [51,52]. The EggNOG annotation indicates the presence of eight genes belonging to the F0F1-ATPase system in both native strains and reference WCFS1 (Table 3). Analysis of NGS variant counts indicated that the UTNGt2 strain harbored more variants of the ATP synthase subunits alpha, beta, delta, gamma, epsilon chain, and subunit b, while UTNGt21A displayed variants in subunits alpha and beta only (Table S4). Both strains share the same SNP type of the *atpA* gene, with a conservative in-frame insertion (c.1501\_1502insCCGCTG; p.Thr501\_Ala502insAlaAla) and an upstream gene multi-nucleotide variant (TCCC). Genome resequencing indicated the presence of several SNPs (synonymous\_variant) of the *lepA\_1* gene (encoding for an elongation factor A), with one missense\_variant detected in UTNGt21A only (c.1354G>T; p.Val452Phe). Furthermore, four missense variants (c.82T>G; p.Ser28Ala; c.850A>G; p.Thr284Ala; c.856C>T; p.Arg286Trp; and c.1150G>A; p.Val384Met) and eight synonymous\_variants within the *lepA\_2* gene were detected in the UTNGt2 genome only (Table 3). It has been found that the elongation factors are involved in the attachment of bacteria to the human intestine and mucins [53]. How these mutations will affect the strains' bacterial cell attachment requires further investigation. Nevertheless, we suggest that the fruit acidic microenvironment heritage and the genetic resistance machinery of these strains might support their further survival in low-pH habitats.

A crucial parameter for the survival of the strain in the digestive tract is bile tolerance [54]. Early research indicates that the bile salt hydrolase gene *bsh* might be responsible for the tolerance of some *L. plantarum* strains to bile [55]. Nevertheless, there is no evidence of a direct connection between Bsh and bile resistance [54]. In the genomes of both native strains, no *bsh* genes were annotated. Instead, a conjugated bile acid hydrolase gene *cbh* (lp\_3536) was detected along with three putative choloylglycine hydrolases (*yxel\_1*, *yxel\_2*, and *yxel\_3*)—proteins involved in bile hydrolysis (Table 3). However, complementary in vitro analysis indicated that both native strains can adapt and maintain for 4 hours in 0.3% bile salt, with a slight increase in viability for UTNGt2 (Figure S5).

The LAB strains are exposed to various temperatures, as the products that incorporate the cells must fulfill the safety requirements of the food sector [56]. Thus, several genes encoding proteins involved in heat shock stress resistance—such as 18kDa, heat shock protein, protein 15, and GrpE (methionine synthase)—were annotated with EggNOG within the genomes of both strains. Previous research indicates that by exposure to acidic pH (2.0) and incubation at 37 °C the GrpE and 30S ribosomal proteins were induced in some *L. plantarum* strains isolated from olive and corn silage [57]. Likewise, the chaperonins (GroEL) and cofactors (GroES) that play essential roles in promoting the correct folding and subsequent translocation of nascent polypeptides were annotated in both native strains (Table 3). Furthermore, two genes—*ccpA* and *ccpB*, encoding a catabolite control proteins A and B, respectively—were annotated in both native strains. The lack of these proteins affected the survival of the strain under heat [58].



**Table 3.** List of genes and annotated variants associated with bacterial adaptation to different stressors in specific microenvironments.

Stress Factor	Gene (locus WCFS1)	% Identity (EggNOG Annotation)/No. of Variants Relative to the REFERENCE WCFS1				
		Protein Product	UTNGt21A		UTNGt2	
pH	<i>atpC</i> (lp_2363)	ATP synthase epsilon chain	67.60	(-)	67.60	1
	<i>atpD</i> (lp_2364)	ATP synthase subunit beta	84.79	2	84.79	1
	<i>atpG</i> (lp_2365)	ATP synthase gamma chain	64.19	(-)	64.19	2
	<i>atpA</i> (lp_2366)	ATP synthase subunit alpha	81.34	3	81.34	2
	<i>atpH</i> (lp_2367)	ATP synthase subunit delta	45.55	(-)	45.55	1
	<i>atpF</i> (lp_2368)	ATP synthase subunit b	57.64	(-)	57.64	1
	<i>atpE</i> (lp_2369)	ATP synthase subunit c	82.69	(-)	82.69	(-)
	<i>atpB</i> (lp_2370)	ATP synthase subunit a	54.85	(-)	54.85	(-)
	<i>lepA_1</i> (lp_2015)	Elongation factor 4	56.47	3	82.75	5
	<i>lepA_2</i> (lp_3120)	Elongation factor 4	82.75	(-)	56.63	12
Bile salt hydrolase	<i>yxel_1</i>	Putative protein YxeI (Choloylglycine hydrolase)	42.98	(-)	42.98	(-)
	<i>yxel_2</i>	Putative protein YxeI (Choloylglycine hydrolase)	40.54	(-)	34.85	(-)
	<i>yxel_3</i>	Putative protein YxeI (Choloylglycine hydrolase)	34.85	(-)	40.55	(-)
	<i>cbh</i> (lp_3536)	Conjugated bile acid hydrolase	67.28	(-)	67.28	(-)
	<i>hsp2</i> (lp_2668)	18 kDa heat shock protein	44.96	3	42.05	3
	<i>hrcA</i> (lp_2029)	Heat-inducible transcription repressor HrcA	58.90	(-)	58.90	(-)
	<i>grpE</i> (lp_2028)	Protein GrpE	58.89	(-)	58.89	1
	<i>dnaK</i> (lp_2027)	Chaperone protein DnaK	84.33	(-)	84.33	3
	<i>dnaJ</i> (lp_2026)	Chaperone protein DnaJ	71.12	3	71.12	3
	<i>Gt21A_00947</i> <i>Gt21A_01250Gt2_02817</i>	18 kDa heat shock protein	44.9633.82	(-)	44.96	(-)
Temperature	<i>hslR</i>	Heat shock protein 15	70.79	(-)	71.91	(-)
	<i>groL</i> (lp_0728)	60 kDa chaperonin	84.89	2	84.89	2
	<i>groS</i> (lp_0727)	10 kDa chaperonin	69.14	(-)	69.14	1
	<i>hslO</i> (lp_0548)	33 kDa chaperonin	69.61	2	69.61	(-)
	<i>hsp 1</i> (lp_0129)	Hypothetical small heat shock protein	45.28	1	45.28	3
	<i>ccpA_1</i>	Catabolite control protein A	49.33		49.33	
	<i>ccpA_2</i>	Catabolite control protein A	44.09	3	44.09	4
	<i>ccpA_3</i>	Catabolite control protein A	65.76		65.76	
	<i>ccpA_4</i>	Catabolite control protein A	(-)		44.14	
	<i>ccpB</i>	Catabolite control protein B	44.92	1	47.60	2
<i>cspP</i> (lp_1160)	Cold shock protein 1	78.78	1	78.78	1	
<i>cspL</i> (lp_0031)	Cold shock protein 2	81.81	(-)	81.81	(-)	
<i>cspLA</i>	Cold shock-like protein CspLA	86.36	(-)	86.36	(-)	
<i>opuCD</i> (lp_1610)	Carnitine transport permease protein OpuCD Glycine	73.15	2	73.17	2	
Osmosis	<i>opuCC</i> (lp_1609)	betaine/carnitine/choline-binding protein OpuCC	63.10	5	63.49	2
	<i>opuCB_1</i> (lp_1608)	Carnitine transport permease protein OpuCB	98.21	5	75.00	3
	<i>opuCA</i> (lp_1607)	Carnitine transport ATP-binding protein OpuCA	68.62	3	68.62	3
	<i>opuCB_2</i>	Carnitine transport permease protein OpuCB Glycine	75.00	(-)	(-)	(-)
	<i>choS</i> (lp_0367)	betaine/carnitine/choline-binding protein Glycine	73.14	6	73.14	24
	<i>choQ</i> (lp_0368)	betaine/carnitine/choline-binding protein	68.62	3	68.62	(-)

(-): No variant detected.

Cold stress resistance is a determining factor of probiotic features [59]. Cold or freezing can diminish the cell viability considerably, inducing cell membrane damage and, consequently, altering the aromatic molecules [60]; thus, to overcome these harmful effects, the LAB genome is equipped with cold stress response genes encoding proteins such as CSPs [61]. Three cold stress proteins (*csp*, *cspL*, and *cspLA*) were annotated in both native strains. One SNP of the *csp* gene (c.144T>C; p.Thr48Thr) was found. In a recent complementary in vitro evaluation of the UTNGt2 cell viability changes during storage, under freezing (−80 °C) conditions, in a juice matrix, a decrease of 10% was detected in the samples stored for 12 months; nevertheless, the number of cells was greater ( $1.28 \times 10^7$  CFU/mL) than the probiotic threshold value ( $1.0 \times 10^6$  CFU/mL), suggesting high adaptability to cold stress (unpublished data). Similar results (5% viability decrease) were obtained after the storage of a milk-based product supplemented with UTNGt2 for 21 days under refrigeration—the remaining cells ( $2.09 \times 10^7$  CFU/mL) were superior to the threshold limit, indicating high adaptability under cold conditions. A different performance was obtained with the UTNGt21A strain, where the cell viability decreased in milk but not in the fruit matrix (unpublished data).

During food processing, the LAB strains must tolerate the osmotic changes that can affect the growth of the cells and their metabolism; therefore, they must overcome these changes by developing a molecular system against osmosis stress that includes two ABC transporters (*opuABCD*, *choSQ*) [57]. Early genomic characterization of the reference *L. plantarum* WCFS1 indicates that these systems are related to the biosynthesis and uptake of the osmoprotectants such as glycine, betaine, carnitine, and choline [9]. A divergent number of variants were annotated. Among them, two missense\_variants of the *opuB* (c.136C>T; p.Pro46Ser) and *opuC* (c.567G>T; p.Lys189Asn) genes were detected in the UTNGt21A genome only (Table 3). An early study indicated that these molecules might protect against salt-induced cell injury [47]. Moreover, a superior number of mutations (19) were annotated within the genes (*mutS*, *mutY*, and *mutT*) encoding for proteins involved in the DNA mismatch repair mechanism of the UTNGt21A genome. In contrast, 10 mutations were detected within the UTNGt2 genome. Previous studies suggest that there is a correlation between the frequency of mutations within the proteins involved in the repair mechanisms and the adaptation of the strain to the dairy matrix [7]. As already mentioned above, at this point, we cannot predict whether these variants are producing a phenotype of interest. Further in vitro analysis might help to understand their achieved probiotic properties in the desired food matrix or their effectiveness as antimicrobial-producing strains.

#### 4. Conclusions

The identification of SNPs from the whole-genome resequencing is challenging in the context of the LAB strains' adaptability to various niches and their further biotechnological performance. In this study, the genomic comparison of two native lactobacilli strains illustrates a distinct genetic repertoire and divergent gene variants that could be inherent to their microenvironment niches. Moreover, genomic comparison revealed that native lactobacilli harbored a plethora of genes participating in antimicrobial activity, with the UTNGt21A strain displaying a more complex bacteriocin cluster gene organization than UTNGt2. Pan-genome analysis disclosed the presence of various strain-specific genes involved in the defense mechanism and carbohydrate metabolism. The divergence between the functional gene cassettes detected within the GIs related to the carbohydrate metabolism might explain the strains' response to designated niches. In addition, both native strains shared a common genetic pattern for stress tolerance; nevertheless, they were divergent in their types of gene variants, which might alter their adaptation and effectiveness if intended to be applied in different food matrices. The SNP counts and classes, the types of detected variants, and the impact of accumulated mutations on the strain phenotypes remain poorly understood, but can provide insights into the dynamics of the adaptation of LAB to different stressors or environments, as well as their probable technological applicability. Further in vitro analysis will direct the evaluation of probiotic technological properties and

antimicrobial strength in different food matrices in order to better understand the impact of these variations on the overall strain utility.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/genes13030443/s1>. Table S1: GIs detected within the (A) UTNGt21A and (B) UTNGt2 genomes; Table S2: Base change count for every SNP of (A) UTNGt21A and (B) UTNGt2; Table S3: Variant annotation results of putative plantaricin genes of native *L. plantarum* UTNGt21A and UTNGt2 strains; Table S4: Variant annotation results of genes related to strains' adaptation to different stressors; Figure S1: (A) Circular and (B) linear representation of the whole-genome comparison of *L. plantarum* UTNGt21A, UTNGt2, and WCFS1; Figure S2: Distribution and type of IS elements annotated with EggNOG and the Issaga web tool; Figure S3: (A) Pairwise protein sequence alignment of putative bacteriocin from UTNGT21A and enterolysin\_A from UTNGt21O and *Enterococcus faecalis* (NCBI accession no. AGG79281.1), with Clustal WS (v.2.0) retrieved from Jalview; (B) average distance calculated based on percentage similarity between sequences; Figure S4: Different expression profiles of *S. aureus* ATCC1026 treated with different peptide extracts; Figure S5: The effect of 0.3% bile on strain viability.

**Author Contributions:** G.N.T. was responsible for conceptualization, methodology, data curation, supervision, project administration, funding acquisition, writing, review, and editing. The author has read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Centre of Research (CUICYT) of the Technical University of the North, Grant No. 2929/2019.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** The author acknowledges the Technical University of the North, Centre of Research (CUICYT), for funding this work, and expresses their sincere gratitude to D. Brown for carefully reviewing and correcting the manuscript.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Sebastianes, F.L.S.; de Azevedo, J.L.; Souza-Motta, C.M. Diversity and biotechnological potential of endophytic microorganisms associated with tropical mangrove forests. In *Diversity and Benefits of Microorganisms from the Tropics*; de Azevedo, J.L., Quecine, M.C., Eds.; Springer: Berlin/Heidelberg, Germany, 2017; pp. 37–57. [\[CrossRef\]](#)
2. Abe Sato, S.T.; Marques, J.M.; da Luz de Freitas, A.; Sanches Progenio, R.C.; Nunes, M.R.T.; Mota de Vasconcelos Mas-safra, J.; Gomes Moura, F.; Rogez, H. Isolation and genetic identification of endophytic lactic acid bacteria from the Amazonian açai fruits: Probiotics features of selected strains and their potential to inhibit pathogens. *Front. Microbiol.* **2021**, *11*, 610524. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Jabbar, Z.; Mukhtar, H.; Tayyeb, A.; Manzoor, A. Next-generation sequencing to elucidate adaptive stress response and plantaricin genes among *Lactobacillus plantarum* strains. *Future Microbiol.* **2020**, *15*, 333–348. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Hernández-González, J.C.; Martínez-Tapia, A.; Lazcano-Hernández, G.; García-Pérez, B.E.; Castrejón-Jiménez, N.S. Bacteriocins from lactic acid bacteria. A powerful alternative as antimicrobials, probiotics, and immunomodulators in veterinary medicine. *Animals* **2021**, *11*, 979. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Douillard, F.P.; Kant, R.; Ritari, J.; Paulin, L.; Palva, A.; de Vos, W.M. Comparative genome analysis of *Lactobacillus casei* strains isolated from Actimel and Yakult products reveals marked similarities and points to a common origin. *Microb. Biotechnol.* **2013**, *6*, 576–587. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Botta, C.; Acquadro, A.; Greppi, A.; Barchi, L.; Bertolino, M.; Cocolin, L.; Rantsiou, K. Genomic assessment in *Lactobacillus plantarum* links the butyrogenic pathway with glutamine metabolism. *Sci. Rep.* **2017**, *7*, 15975. [\[CrossRef\]](#)
7. Bachmann, H.; Starrenburg, M.J.; Molenaar, D.; Kleerebezem, M.; van Hylckama Vlieg, J.E.T. Microbial domestication signatures of *Lactococcus lactis* can be reproduced by experimental evolution. *Genome Res.* **2012**, *22*, 115–124. [\[CrossRef\]](#)
8. Manasian, P.; Bustos, A.S.; Pålsson, B.; Håkansson, A.; Peñarrieta, J.M.; Nilsson, L.; Linares-Pastén, J.A. First evidence of acyl-hydrolase/lipase activity from human probiotic bacteria: *Lactobacillus rhamnosus*GG and *Bifidobacterium longum* NCC 2705. *Front. Microbiol.* **2020**, *11*, 1534. [\[CrossRef\]](#)
9. Kleerebezem, M.; Boekhorst, J.; van Kranenburg, R.; Molenaar, D.; Kuipers, O.P.; Leer, R.; Tarchini, R.; Peters, S.A.; Sandbrink, H.M.; Fiers, M.W.; et al. Complete genome sequence of *Lactobacillus plantarum* WCFS1. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1990–1995. [\[CrossRef\]](#)

10. De Vries, M.C.; Vaughan, E.E.; Kleerebezem, M.; de Vos, W.M. *Lactobacillus plantarum*—Survival, functional and potential probiotic properties in the human intestinal tract. *Int. Dairy J.* **2006**, *16*, 1018–1028. [[CrossRef](#)]
11. 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](#)]
12. Choi, S.; Jin, G.D.; Park, J.; You, I.; Kim, E.B. Pan-Genomics of *Lactobacillus plantarum* revealed group-specific genomic profiles without habitat association. *J. Microbiol. Biotechnol.* **2018**, *28*, 1352–1359. [[CrossRef](#)]
13. Mao, B.; Yin, R.; Li, X.; Cui, S.; Zhang, H.; Zhao, J.; Chen, W. Comparative genomic analysis of *Lactiplantibacillus plantarum* isolated from different niches. *Genes* **2021**, *12*, 241. [[CrossRef](#)]
14. Barraclough, T.; Balbi, K.; Ellis, R. Evolving concepts of bacterial species. *Evol. Biol.* **2012**, *39*, 148–157. [[CrossRef](#)]
15. Naeem, M.; Ilyas, M.; Haider, S.; Baig, S.; Saleem, M. Isolation characterization and identification of lactic acid bacteria from fruit juices and their efficacy against antibiotics. *Pak. J. Bot.* **2012**, *44*, 323–328.
16. Di Cagno, R.; Filannino, P.; Gobbetti, M. Vegetable and fruit fermentation by lactic acid bacteria. In *Biotechnology of Lactic Acid Bacteria: Novel Applications*; Mozzi, F., Raya, R.R., Vignolo, G.M., Eds.; John Wiley & Sons: Chichester, UK, 2015; pp. 216–230. [[CrossRef](#)]
17. Benavidez, A.; Ulcuango, M.; Yépez, L.; Tenea, G.N. Assessment of the in vitro bioactive properties of lactic acid bacteria isolated from native ecological niches of Ecuador. *Rev. Argent. Microbiol.* **2016**, *48*, 236–244. [[CrossRef](#)] [[PubMed](#)]
18. Tenea, G.N.; Jarrin-V, P.; Yepez, L. Microbiota of wild fruits from the amazon region of Ecuador: Linking diversity and functional potential of lactic acid bacteria with their origin. In *Ecosystem and Biodiversity of Amazonia*; Mikkola, H.J., Ed.; IntechOpen: London, UK, 2020. [[CrossRef](#)]
19. Tenea, G.N.; Ortega, C. Genome characterization of *Lactiplantibacillus plantarum* strain UTNGt2 originated from *Theobroma grandiflorum* (white cacao) of Ecuadorian Amazon: Antimicrobial peptides from safety to potential applications. *Antibiotics* **2021**, *10*, 383. [[CrossRef](#)]
20. Tenea, G.N.; Hurtado, P.; Ortega, C. A novel *Weissella cibaria* strain UTNGt21O isolated from wild *Solanum quitoense* Fruit: Genome sequence and characterization of a peptide with highly inhibitory potential toward gram-negative bacteria. *Foods* **2020**, *9*, 1242. [[CrossRef](#)]
21. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)]
22. Grant, J.R.; Stothard, P. The CGView Server: A comparative genomics tool for circular genomes. *Nucleic Acids Res.* **2008**, *36*, W181–W184. [[CrossRef](#)] [[PubMed](#)]
23. Darling, A.C.; Mau, B.; Blattner, F.R.; Perna, N.T. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* **2004**, *14*, 1394–1403. [[CrossRef](#)]
24. 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](#)]
25. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)] [[PubMed](#)]
26. Bertelli, C.; Laird, M.R.; Williams, K.P.; Simon Fraser University Research Computing Group; Lau, B.Y.; Hoad, G.; Winsor, G.L.; Brinkman, F. IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* **2017**, *45*, W30–W35. [[CrossRef](#)] [[PubMed](#)]
27. Siguier, P.; Perochon, J.; Lestrade, L.; Mahillon, J.; Chandler, M. ISfinder: The reference centre for bacterial insertion sequences. *Nucleic Acids Res.* **2006**, *34*, D32–D36. [[CrossRef](#)] [[PubMed](#)]
28. Tarasov, A.; Vilella, A.J.; Cuppen, E.; Nijman, I.J.; Prins, P. Sambamba: Fast processing of NGS alignment formats. *Bioinformatics* **2015**, *31*, 2032–2034. [[CrossRef](#)] [[PubMed](#)]
29. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
30. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **2011**, *27*, 2987–2993. [[CrossRef](#)]
31. Cingolani, P.; Patel, V.M.; Coon, M.; Nguyen, T.; Land, S.J.; Ruden, D.M.; Lu, X. Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift. *Front. Genetics* **2012**, *3*, 35. [[CrossRef](#)]
32. De Jong, A.; van Hijum, S.A.; Bijlsma, J.J.; Kok, J.; Kuipers, O.P. BAGEL: A web-based bacteriocin genome mining tool. *Nucleic Acids Res.* **2006**, *34*, W273–W279. [[CrossRef](#)]
33. Huang, T.; Xiong, T.; Peng, Z.; Xiao, Y.S.; Liu, Z.G.; Hu, M.; Xie, M.Y. Genomic analysis revealed adaptive mechanism to plant-related fermentation of *Lactobacillus plantarum* NCU116 and *Lactobacillus* spp. *Genomics* **2020**, *112*, 703–711. [[CrossRef](#)]
34. Turner, M.S.; Hafner, L.M.; Walsh, T.; Giffard, P.M. Identification, characterization and specificity of a cell wall lytic enzyme from *Lactobacillus fermentum* BR11. *FEMS Microbiol. Lett.* **2004**, *238*, 9–15. [[CrossRef](#)] [[PubMed](#)]
35. Li, P.; Li, X.; Gu, Q.; Lou, X.-Y.; Zhang, Z.-M.; Song, D.-F.; Zhang, C. Comparative genome analysis of *Lactobacillus plantarum* ZJ316 reveals its genetic adaptation and potential probiotic profiles. *J. Zhejiang Univ. Sci. B* **2016**, *18*, 569–579. [[CrossRef](#)]
36. 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](#)]



37. Burgess, C.; O'connell-Motherway, M.; Sybesma, W.; Hugenholtz, J.; van Sinderen, D. Riboflavin production in *Lactococcus lactis*: Potential for in situ production of vitamin-enriched foods. *Appl. Environ. Microbiol.* **2004**, *70*, 5769–5777. [[CrossRef](#)]
38. Alan, Y. Culture fermentation of *Lactobacillus* in traditional pickled gherkins: Microbial development, chemical, biogenic amine and metabolite analysis. *J. Food. Sci. Technol.* **2019**, *56*, 3930–3939. [[CrossRef](#)]
39. Cai, H.; Thompson, R.; Budinich, M.F.; Broadbent, J.R.; Steele, J.L. Genome sequence and comparative genome analysis of *Lactobacillus casei*: Insights into their niche-associated evolution. *Genome Biol. Evol.* **2009**, *1*, 239–257. [[CrossRef](#)]
40. Molenaar, D.; Bringel, F.; Schuren, F.H.; de Vos, W.M.; Siezen, R.J.; Kleerebezem, M. Exploring *Lactobacillus plantarum* genome diversity by using microarrays. *J. Bacteriol.* **2005**, *187*, 6119–6127. [[CrossRef](#)]
41. Wang, J.; Raskin, L.; Samuels, D.C.; Shyr, Y.; Guo, Y. Genome measures used for quality control are dependent on gene function and ancestry. *Bioinformatics* **2015**, *31*, 318–323. [[CrossRef](#)] [[PubMed](#)]
42. Deveson, I.W.; Brunck, M.E.; Blackburn, J.; Tseng, E.; Hon, T.; Clark, T.A.; Clark, M.B.; Crawford, J.; Dinger, M.E.; Nielsen, L.K.; et al. Universal alternative splicing of noncoding exons. *Cell Syst.* **2018**, *6*, 245–255.e5. [[CrossRef](#)] [[PubMed](#)]
43. Nesmelova, I.V.; Hackett, P.B. DDE transposases: Structural similarity and diversity. *Adv. Drug Deliv. Rev.* **2010**, *62*, 1187–1195. [[CrossRef](#)]
44. Smeets, L.C.; Bijlsma, J.J.; Kuipers, E.J.; Vandenbroucke-Grauls, C.M.; Kusters, J.G. The *dprA* gene is required for natural transformation of *Helicobacter pylori*. *FEMS Immunol. Med. Microbiol.* **2000**, *27*, 99–102. [[CrossRef](#)]
45. Zhang, B.; Zuo, F.; Yu, R.; Zeng, Z.; Ma, H.; Chen, S. Comparative genome-based identification of a cell wall-anchored protein from *Lactobacillus plantarum* increases adhesion of *Lactococcus lactis* to human epithelial cells. *Sci. Rep.* **2015**, *5*, 14109. [[CrossRef](#)] [[PubMed](#)]
46. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.; Clamp, M.; Barton, G.J. Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191. [[CrossRef](#)]
47. Tenea, G.N. Peptide extracts from native lactic acid bacteria generate ghost cells and spheroplasts upon interaction with *Salmonella enterica*, as promising food antimicrobials. *BioMed Res. Int.* **2020**, *2020*, 6152356. [[CrossRef](#)]
48. Mbye, M.; Baig, M.A.; AbuQamar, S.F.; El-Tarabily, K.A.; Obaid, R.S.; Osaili, T.M.; Al-Nabulsi, A.A.; Turner, M.S.; Shah, N.P.; Ayyash, M.M. Updates on understanding of probiotic lactic acid bacteria responses to environmental stresses and highlights on proteomic analyses. *Compr. Rev. Food Sci. Food Saf.* **2020**, *19*, 1110–1124. [[CrossRef](#)] [[PubMed](#)]
49. EFSA FEEDAP Panel. Guidance on the characterization of microorganisms used as feed additives or as production organisms. *EFSA J.* **2018**, *16*, 5206. [[CrossRef](#)]
50. De Angelis, M.; Gobbetti, M. Environmental stress responses in *Lactobacillus*: A review. *Proteomics* **2004**, *4*, 106–122. [[CrossRef](#)] [[PubMed](#)]
51. Haddaji, N.; Krifi, B.; Lagha, R.; Khouadja, S.; Bakhrouf, A. Effect of high temperature on viability of *Lactobacillus casei* and analysis of secreted and GroEL proteins profiles. *Afr. J. Bacteriol. Res.* **2015**, *7*, 29–35. [[CrossRef](#)]
52. Jiang, Y.; Zhang, J.; Zhao, X.; Zhao, W.; Yu, Z.; Chen, C.; Yang, Z. 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.* **2018**, *82*, 1225–1233. [[CrossRef](#)]
53. Altermann, E.; Russell, W.M.; Azcarate-Peril, M.A.; Barrangou, R.; Buck, B.L.; McAuliffe, O.; Souther, N.; Dobson, A.; Duong, T.; Callanan, M.; et al. Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 3906–3912. [[CrossRef](#)]
54. Lambert, J.M.; Bongers, R.S.; de Vos, W.M.; Kleerebezem, M. Functional analysis of four bile salt hydrolase and penicillin acylase family members in *Lactobacillus plantarum* WCFS1. *Appl. Environ. Microbiol.* **2008**, *74*, 4719–4726. [[CrossRef](#)] [[PubMed](#)]
55. Fang, F.; Li, Y.; Bumann, M.; Raftis, E.J.; Casey, P.G.; Cooney, J.C.; Walsh, M.A.; O'Toole, P.W. Allelic variation of bile salt hydrolase genes in *Lactobacillus salivarius* does not determine bile resistance levels. *J. Bacteriol.* **2009**, *191*, 5743–5757. [[CrossRef](#)] [[PubMed](#)]
56. Hamon, E.; Horvatovich, P.; Izquierdo, E.; Bringel, F.; Marchioni, E.; Aoudé-Werner, D.; Ennahar, S. Comparative proteomic analysis of *Lactobacillus plantarum* for the identification of key proteins in bile tolerance. *BMC Microbiol.* **2011**, *11*, 63. [[CrossRef](#)]
57. Castaldo, C.; Siciliano, R.A.; Muscariello, L.; Marasco, R.; Sacco, M. CcpA affects expression of the groESL and dnaK operons in *Lactobacillus plantarum*. *Micro. Cell Fact.* **2006**, *5*, 35. [[CrossRef](#)] [[PubMed](#)]
58. Lorv, J.S.; Rose, D.R.; Glick, B.R. Bacterial ice crystal controlling proteins. *Scientifica* **2014**, *2014*, 976895. [[CrossRef](#)]
59. Song, S.; Bae, D.-W.; Lim, K.; Griffiths, M.W.; Oh, S. Cold stress improves the ability of *Lactobacillus plantarum* 167 to survive freezing. *Int. J. Food Microbiol.* **2014**, *191*, 135–143. [[CrossRef](#)]
60. Polo, L.; Manes-Lazaro, R.; Olmeda, I.; Cruz-Pio, L.E.; Medina, A.; Ferrer, S.; Pardo, I. Influence of freezing temperatures prior to freeze-drying on viability of yeasts and lactic acid bacteria isolated from wine. *J. Appl. Microbiol.* **2017**, *122*, 1603–1614. [[CrossRef](#)]
61. Wood, J.M. Bacterial responses to osmotic challenges. *J. Gen. Physiol.* **2015**, *145*, 381–388. [[CrossRef](#)]