**In silico** genomic analysis of the potential probiotic *Lactiplantibacillus pentosus* CF2-10N reveals promising beneficial effects with health promoting properties

Hikmate Abriouel*, Julia Manetsberger, Natacha Caballero Gómez and Nabil Benomar

Área de Microbiología, Departamento de Ciencias de la Salud, Facultad de Ciencias Experimentales, Universidad de Jaén, Jaén, Spain

*Lactiplantibacillus pentosus* CF2-10N, isolated from brines of naturally fermented Aloreña green table olives, exhibited high probiotic potential. High throughput sequencing and annotation of genome sequences underline the potential of *L. pentosus* CF2-10N as excellent probiotic candidate of vegetable origin. In a previous study we could show the probiotic potential of CF2-10N *in vitro*, while in this study *in silico* analysis of its genome revealed new insights into its safety and functionality. Our findings highlight the microorganism's ecological flexibility and adaptability to a broad range of environmental niches, food matrices and the gastrointestinal tract. These features are shared by both phylogenetically very close *L. pentosus* strains (CF2-10N and MP-10) isolated from the same ecological niche with respect to their genome size (≅3.6 Mbp), the presence of plasmids (4–5) and several other properties. Nonetheless, additional and unique features are reported in the present study for *L. pentosus* CF2-10N. Notably, the safety of *L. pentosus* CF2-10N was shown by the absence of virulence determinants and the determination of acquired antibiotic resistance genes, i.e., resistome, which is mostly represented by efflux-pump resistance genes responsible for the intrinsic resistance. On the other hand, defense mechanisms of *L. pentosus* CF2-10N include eight prophage regions and a CRISPR/cas system (CRISPR-I and CRISPR-II) as acquired immune system against mobile elements. Finally, the probiotic potential of this strain was further demonstrated by the presence of genes coding for proteins involved in adhesion, exopolysaccharide biosynthesis, tolerance to low pH and bile salts, immunomodulation, and vitamin and enzyme production. Taken together these results, we propose the use of *L. pentosus* CF2-10N as a potential and promising probiotic candidate able to colonize several niches and adapt to different lifestyles. The strain can provide attractive functional and probiotic features necessary for its application as starter culture and probiotic.

**KEYWORDS**
Aloreña table olives, *Lactiplantibacillus pentosus*, probiotics, *in silico* analysis, safety, functional properties

---

**Funding**
This study was supported by the University of Jaén (Jaén, Spain) through the Project: Genomic characterization of the probiotic *Lactiplantibacillus pentosus* CF2-10N: safety, functional properties and ecological flexibility, grant number: 21202/2018.

**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.

**Consent for publication**
The authors declare that the submission contains no patient data.
Introduction

Probiotics are defined by the Food and Agriculture Organization of the United Nations/World Health Organization (FAO/WHO) “as live microorganisms that, when administered in adequate amounts, confer a health benefit to the host” (Hill et al., 2014). In this regard, probiotic microorganisms are characterized by their diverse origin, taxonomy, fitness, effective dose, host and health benefits depending specifically on the strain employed. Thus, preliminary screening criteria for potential probiotic microorganisms include their capacity to withstand several barriers and challenges (1) in vitro, such as stressful environmental conditions; and (2) in vivo—notably during their passage through the gastrointestinal tract (acids and bile salts), their capacity to adhere and colonize human epithelial cells and their ability to produce beneficial effects in the host (antimicrobial activity, modulation of the immune system, degradation of toxic components, etc.).

In this sense, the key element for the differentiation of probiotic strains from each other is their specific functionality. Naturally, this has led to a considerable amount of research efforts put into determining the specific probiotic effect(s) of each potential probiotic strain and highlighting their potential targets over recent years (Allain et al., 2018; van de Wijgert and Verwijs, 2019; Yan et al., 2019; Yoha et al., 2022). In other words, a search for unique and attractive functional characteristics is crucial to provide new and helpful information on microorganisms with probiotic potential. This is especially important for those microorganisms that are naturally present in fermented foods, such as for example Lactiplantibacillus species.

On the other hand, probiotics as indicated by their name, act as a ‘promoter of life’ supporting in a natural way the improvement of the overall health status of the host organism (Amara and Shibil, 2015). It has further been shown that it is possible to combine several of these strains into multi-strain probiotics (Nayak, 2010), where the strains of this ‘probiotic cocktail’ can work synergistically, thus greatly increasing the overall benefit spectrum for the host (Puvanasundram et al., 2021).

The recently reclassified Lactiplantibacillus pentosus, formerly known as Lactobacillus pentosus (Zheng et al., 2020), colonizes a large set of environmental niches and therefore exhibits a huge ecological and metabolic adaptability (Anukam et al., 2013; Abriouel et al., 2017; Pérez-Díaz et al., 2021). Due to its genomic diversity and functionality, this species is found in several fermented foods (vegetables, meat, and dairy), plants, animals, vaginal, urogenital and gastrointestinal tract, while also having a large set of biotechnological and probiotic applications (Tofalo et al., 2014; Vaccalluzzo et al., 2020).

Lactiplantibacillus pentosus together with L. plantarum is an important member of the bacterial community found on the surface of olive fruits and thus represent the predominating bacteria in olive fermentation. Notably, they promote the fermentation process, conservation and extension of shelf life of the product, in addition to their role in organoleptic properties and the production of health promoting molecules such as amino acids, short chain fatty acids (SCFA), antioxidatns, exopolysaccharides and vitamins (Caggianiello et al., 2016; Carrasco et al., 2018; Benitez-Cabello et al., 2019; Perpetuini et al., 2020). Furthermore, besides the production of the abovementioned molecules in foods such as olives, these bacteria are also able to produce these substances in vivo, i.e., in the gastrointestinal tract thus providing an important probiotic effect (Oguntoyinbo and Narbad, 2015; Saxami et al., 2017; Gualitario et al., 2018). Consequently, several fermented foods have been classified as functional foods, as they are carriers of probiotic organisms and/or their molecules. In this regard, the health benefit and functionality of table olives goes beyond just ‘fermented food’ due to their ability to deliver beneficial microbes adhering to the drupe epidermis into the human gastrointestinal tract where they may influence the microbial diversity and functionality (Lavermicocca et al., 2005; Rodríguez-Gómez et al., 2014, 2017).

Among olives, naturally fermented Aloreña green table olives are a promising carrier of probiotics since they are characterized by their diverse microbial community. This is mostly due to the richness of the ecosystem (soil, plant, and brine) and the progressive changes inherent to the production process (Abriouel et al., 2011).

The microbial diversity of Aloreña table olives includes lactic acid bacteria (LAB), mainly L. pentosus-yeasts and other contaminant microorganism, with microbial profiles greatly depending on the fermentation conditions (e.g., vat, fermenter or in cold). In this regard, however under vat and fermenter conditions, LAB and yeasts have been determined as the main actors (Abriouel et al., 2011). Among LAB, L. pentosus are considered potential probiotics due to their good growth capacity and survival rate under simulated gastro-intestinal conditions (acidic pH of 1.5, up to 4% of bile salts and 5 mM of nitrate), auto-aggregation, co-aggregation with pathogenic bacteria, adhesion to intestinal and vaginal cell lines, biofilm formation, fermentation of several prebiotics and their capacity to ferment lactose among others (Pérez Montoro et al., 2016). In addition, omics approaches were used by our group; including genomics, proteomics and transcriptomics, to determine and confirm the safety and functionality of the probiotic L. pentosus isolated from Aloreña table olives (Casado Muñoz et al., 2016; Pérez Montoro et al., 2018a, b; Alonso García et al., 2021, 2023).

Hence, in the present study, we extend the characterization of L. pentosus using in silico genomic analysis to unveil the genetic basis of the safety and probiotic ability of L. pentosus CF2-10N – one of the most promising potential probiotic strains isolated from Aloreña table olives (Abriouel et al., 2012).
Materials and methods

Bacterial strain and growth conditions

*Lactiplantibacillus pentosus* CF2-10 N, originally isolated from naturally fermented Aloreña green table olives (Abriouel et al., 2012), was selected based on its probiotic profile as reported by Pérez Montoro et al. (2016). *Lactiplantibacillus pentosus* CF2-10 N was routinely cultured at 37°C in de Man, Rogosa and Sharpe (MRS) broth or agar (Fluka, Madrid, Spain) under aerobic (atmospheric) conditions for 24–48 h. The strain was kept in 20% glycerol at −80°C for long-term storage.

DNA extraction, library preparation and genome sequencing

Bacterial cells of *L. pentosus* CF2-10 N were harvested by centrifugation after 18 h incubation at 37°C under aerobic conditions in liquid medium. Total genomic DNA was obtained using the PureGene core kit B, according to the manufacturer's instructions (Qiagen, Spain). DNA quantification and quality assessment were carried out using a NanoDrop 2000 spectrophotometer (Thermo Scientific), the PicoGreen ds DNA Reagent (Invitrogen) and/or agarose gel electrophoresis (0.8% agarose gel in Tris-borate-EDTA buffer, 90V, 45 min). Bacterial DNA was stored at −20°C until required.

Purified genomic DNA was sheared into 10- to 20-kb fragments using the protocol designed for DNA library preparation using the PacBio RS II System (Pacific Biosciences, Menlo Park, CA, United States). Resulting libraries (22–24 kb) were purified and sequenced using a P6-C4 DNA polymerase (Pacific Biosciences) and single-molecule real-time (SMRT) cells with a 240-min sequence capture protocol and Stage Start to maximize the subread length on the PacBio RS II.

Genome assembly and annotation

Raw sequence data were filtered (Q20) and a total of 150,292 reads were obtained with a median length of 14,991 bp. The resulting reads were assembled *de novo* following the Hierarchical Genome Assembly Process (HGAP3.0) approach (SMRT analysis version: 2.3.0, patch #4) for Pacific Bioscience using the WGS-Celera Assembler 7.0 (Myers et al., 2000) and Quiver algorithm (Chen-Shan Chin et al., 2013). Once assembled, the prediction of Coding DNA Sequences (CDS) was done with the help of the GenMark program (Besemer et al., 2001). Furthermore, prediction of tRNA, rRNA, and mRNA genes and signal peptides in the sequences was achieved using *tRNAscan* (version 2.0), *RNAmmer* (Version 1.2), *HMMer* [HMMER 3.1 (July 2017)]

Comparative genomic analysis of *Lactiplantibacillus pentosus* CF2-10N and other *Lactiplantibacillus pentosus* strains

Genome sequences of *L. pentosus* CF2-10 N and other *L. pentosus* strains (MP-10, IG1 and KCA1) were aligned using MAUVE (Darling et al., 2004) available in DNASTAR Lasergene (version 17.3). Trees were then generated using RAxML with default parameters (Stamatakis, 2014). Further genome alignment and comparison of *L. pentosus* CF2-10 N and other *L. pentosus* strains (IG1 and KCA1) isolated from different ecological niches or *L. plantarum* WCFS1 (as reference strain) was done using the **MUMmer** program (version 3.0), considering alignment > 500 bp. The genome accession numbers of strains used in this study are as follows: *L. pentosus* IG1 (PRJEA67801), *L. pentosus* KCA1 (PRJNA81575, GenBank assembly accession GCA_000203855.3). Functional annotation of CDS (COG) for the three strains (*L. pentosus* IG1, *L. pentosus* KCA1 and *L. plantarum* WCFS1) was completed following the same strategy as for *L. pentosus* CF2-10 N by using reciprocal blast (BLAST2go) program version 4.1.9 (Conesa et al., 2005) and the available genome sequences in NCBI.

Genomic analysis of safety aspects and defense mechanisms of *Lactiplantibacillus pentosus* CF2-10N

For specific annotation of antibiotic resistance genes (ARGs), the Resistance Gene Identifier (RGI) software (as part of the CARD "The Comprehensive Antibiotic Resistance Database" tools; Alcock et al., 2020) was used for the prediction
of the *L. pentosus* CF2-10 N resistome from protein or nucleotide data based on homology and SNP (Single Nucleotide Polymorphism) models, employing the CARD’s curated AMR (antimicrobial resistance) detection models (last accessed in March 2022). In addition, the genome of *L. pentosus* CF2-10 N was investigated for acquired antibiotic resistance genes/chromosomal mutations mediating antimicrobial resistance through the ResFinder3 software version 4.1 (Zankari et al., 2012; Bortolaia et al., 2020) with selected %ID threshold of 90.00% and selected minimum length of 60% (last accessed in March 2022).

Regarding virulence factors (VFIs), the predicted CDSs were annotated using reciprocal BLAST against the Virulence Factors of Bacterial Pathogens (VFDB) database. Hits were considered positive when the results of reciprocal BLAST were similar, employing a 80% sequence similarity cut-off (Liu et al., 2019).

Concerning mobile genetic elements, the annotated genome sequence of *L. pentosus* CF2-10 N was screened for the presence of conjugative plasmid, transposase, transposon, IS elements and prophage coding genes. The genome was searched for Insertion Sequences (IS) using the ISfinder search tool (Zhang et al., 2000). Furthermore, complementary information on prophage DNA within the *L. pentosus* CF2-10 N genome was obtained by using bioinformatic tools such as PHASTER’s version (PHAge Search Tool Enhanced Release, last updated March 2016; corresponding to the updated prophage/virus database PHAST "PHAge Search Tool") for the rapid identification and annotation of prophage sequences within bacterial genomes and plasmids (Zhou et al., 2011; Arndt et al., 2016).

Finally, the annotated genome sequence of *L. pentosus* CF2-10 N was screened for the presence of CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) coding genes and the localization of CRISPR RNAs targets was identified using the CRISPRDetect program version 2.4 (Biswa et al., 2016).4

### Results

**General genomic features of a probiotic *Lactiplantibacillus pentosus* CF2-10N**

The analysis revealed that the *Lactiplantibacillus pentosus* CF2-10 N genome consisted of a single circular chromosome of 3,645,747 bp, with an estimated mol% G + C content of 46.42% and 4 plasmids ranging 58–120 kb (Figure 1). The annotated genome sequence (Figure 1) revealed 3,713 open reading frames (ORFs), of which 75.4% (2,801) were attributed to a COG (Cluster of Orthologous Groups) family and/or were given a functional description (Supplementary Table S1). Furthermore, 16 rRNA genes were predicted in *L. pentosus* CF2-10 N genome using RNAmmer (version 1.2), while 67 tRNA encoding sequences were identified corresponding to all 20 amino acids and three underlined amino acids (Supplementary Table S2).

Supplementary Figure S1 shows the biological processes, the cellular components and the molecular function frequencies predicted in *L. pentosus* CF2-10 N. Among the Gene Ontology (GO) terms, those related to biological processes such as oxidation–reduction process, regulation of transcription, DNA-templated transcription and DNA-templated transmembrane transport were the most identified. Regarding molecular function, ATP-binding and DNA binding were the most prevalent. However, in both biological process and molecular function about 1,250–1,550 genes have no known biological process/function (Supplementary Figure S1).

The most abundant COG category of *L. pentosus* CF2-10 N genome, except for “[S] Function unknown” (273 CDSs, 9.7%), was “[R] General function prediction only” (336 CDSs, 12%), followed by “[G] Carbohydrate transport and metabolism” (307 CDSs, 11%), “[K] Transcription” (235 CDSs, 8.4%), “[L] Replication, recombination and repair” (213 CDSs, 7.6%) and “[E] Amino acid transport and metabolism” (192 CDSs, 6.9%), accounting for 45.9% of the overall CDS (1,283/2,801 CDSs; Supplementary Table S3).

### Comparative genome analysis of *Lactiplantibacillus pentosus* CF2-10N

Comparative genomic analysis of *L. pentosus* CF2-10N and *L. pentosus* MP-10 isolated from the same ecological niche (Aloreña table olives) showed that both *L. pentosus* strains shared 99.87% identity as revealed by sequence alignment using the MAUVE algorithm. This high similarity was further highlighted by large blocks of colinearization in the MAUVE alignment, being the synteny of genes similar, although inversion, insertion and rearrangement occurred (Figure 2). Besides *L. pentosus* MP-10 (isolated from Aloreña table olives), comparison with other *L. pentosus* strains by genome sequence alignment (using MAUVE), notably IG1 (isolated from olives) and KCA1 (isolated from the vaginal tract), revealed genetic differences among the studied strains (Supplementary Figure S2A). To illustrate this relationship, a maximum-likelihood core genome tree was
constructed using RaxML which showed higher phylogenetic similarity in the case of *L. pentosus* CF2-10 N and MP-10 strains (evolutionary distance “ED,” ED = 0), followed by *L. pentosus* IG1 (ED = 0.02) and then *L. pentosus* KCA1 (ED = 0.08; Supplementary Figure S2B).

The synteny linkage of *L. pentosus* CF2-10 N against *L. pentosus* IG1 and KCA1 strains or *L. plantarum* WCFS1 was further analyzed using the MUMmer program and represented using Circos (Figure 3; Supplementary Tables S4–S6). Here, the genome comparison revealed the presence of highly conserved syntenic blocks between *L. pentosus* strains (IG1 and KCA1; Figures 3A,B), and to a lesser extent with *L. plantarum* WCFS1 (Figure 4C).

In silico analysis of safety determinants and defense mechanisms of *Lactiplantibacillus pentosus* CF2-10N

Safety properties are a crucial feature of potential probiotic strains and their determination is considered a priority when characterizing a new potential probiotic. Hence, in a first step, antibiotic resistance and virulence determinants were screened in the *L. pentosus* CF2-10 N genome sequence. To do so, in silico prediction of antibiotic resistance genes (ARG) was done against the Comprehensive Antibiotic Resistance Database (CARD) using the RGI tool v3.2.1 available in the CARD database 5 which used archive’s curated AMR (antimicrobial resistance) detection models. Results indicated no ARG in the *L. pentosus* CF2-10 N genome sequence. However, BLAST2go annotation revealed the presence of non-specific antimicrobial resistance mechanisms relying on efflux transporters or transmembrane proteins involved in response to antibiotics such as ABC transporter ATP-binding protein (encoded by *LPE_03051*, *LPE_00789*, FD24_GL000501 genes), TIGR00374 family protein (encoded by *mprF* gene), undecaprenyl-diphosphatase (encoded by *uppP* gene), QacE family quaternary ammonium compound efflux SMR transporter (encoded by FD24_GL003284 gene), MATE family efflux transporter (encoded by *LPE_00986* gene) and cation efflux pump (encoded by FD24_GL002035 gene).

With regard to acquired resistance by horizontal gene transfer, ResFinder did not detect any acquired antibiotic resistance genes for aminoglycoside, beta-lactam, colistin, disinfectant, fluoroquinolone, fosfomycin, fusidic acid, glycopeptide, MLS-series (Macrolide, lincosamide and streptogramin B), nitroimidazole, oxazolidinone, phenicol, pseudomonic acid, rifampicin, sulphonamide, tetracycline and trimethoprim (data not shown).

Regarding virulence, the predicted CDSs annotated using reciprocal BLAST against VFDB (database including only experimentally validated virulence factors) did not identify any known virulence factors including toxins.

Analysis of the *L. pentosus* CF2-10 N mobilome showed that the bacterial genome included 66 transposases: 19 transposases, 1 transposase family protein A and 46 transposases belonging to

---

5 https://card.mcmaster.ca/analyze/rgi
nine IS transposase families (4 IS3, 6 IS5, 5 IS21, 17 IS30, 4 IS66, 3 IS1380, 2 IS13, 2 DDE, 2 IS6501, 1 IS200/IS605), mainly located on plasmids (pLPE10-1, pLPE10-2 and pLPE10-4) rather than on the chromosome (50 on plasmids/16 on chromosome) and appearing in multiple copies ranging from two to five (Table 1). IS30 family transposases were abundant (17 of 66 transposases) and were represented by seven different genes (Table 1). Furthermore, Blastp alignment of transposase protein sequences detected in \textit{L. pentosus} CF2-10N genome showed high similarity with \textit{L. pentosus} (29 of 66 transposases, 98.9–100%), \textit{L. plantarum} (11 of 66 transposases, 95.2–100%) and other lactobacilli. It is noteworthy to indicate the presence of 34 paired (adjacent to each other in the genome) transposase genes (2 or 3 genes) being different genes or belonging to different families and located on both chromosome and plasmids (Table 1). Regarding IS elements, 45 CDS were predicted distributed into 16 different families and in various bacteria (Table 2). Here, IS30 and IS3 were the most detected elements followed by IS5 (Table 2).

On the other hand, screening for prophage DNA within the \textit{L. pentosus} CF2-10N genome, using bioinformatic tools such as PHASTER, determined the presence of eight temperate phage regions (Table 3). Two regions were intact (Regions 2 and 3, score > 90), the other three were questionable (Regions 5, 6 and 7,
score 70 ± 90), and the last three regions were incomplete (Regions 1, 4 and 8, score < 70). The complete prophage regions of the L. pentosus CF2-10N chromosome were identified as Lactobacillus phage Sha1 (Regions 2 and 3; GC content, 41.55–41.88%; region length, 39.9–47.7 kb). Regarding the questionable prophage regions, they corresponded to Staphylococcus phage SP beta-like (Regions 5 and 6; GC content, 34.83–40.70%; region length, 13.7–19.4 kb) and Escherichia phage 500,465–1 (Region 7; GC content, 41.54%; region length, 18.8 kb). With respect to the incomplete prophage region, we identified three regions corresponding to Lactobacillus phage PLE3 (Region 1; GC content, 41.26%; region length, 15 kb), Enterobacteria phage fAA91-ss (Region 4; GC content, 38.27%; region length, 23.4 kb) and Escherichia phage 500,465–1 (Region 8; GC content, 31.68%; region length, 6.7 kb; Table 3).

Among the defense mechanisms revealed by in silico analysis of the L. pentosus CF2-10 N genome sequence, CRISPR I and II systems (both signature genes for the Type I "cas3" and Type II "cas9" systems) were detected as defense response to mobile genetic elements (i.e., viruses, transposable elements and conjugative plasmids; Table 4). In this sense, 13 genes were identified as CRISPR associated protein responsible genes (cas genes) organized in two operons (Supplementary Figure S3), and six of them were new genes found in the L. pentosus CF2-10 N genome (Table 4). Regarding CRISPR arrays (CR), five CRISPR unquestionable arrays were identified by using the CRISPRDetect program and they are distributed throughout the genome sequence between 1,791,840 and 3,235,959 bp (Table 5).
Identification of genes associated with probiotic characteristics in Lactiplantibacillus pentosus CF2-10N

In silico genome analysis of probiotic characteristics of L. pentosus CF2-10N revealed the presence of genes coding for adhesion, exopolysaccharide biosynthesis, tolerance to low pH and bile salts, vitamin and enzyme production and immunomodulation among others (Table 6). With respect to adhesion, several genes were identified such as 3 mucus-binding proteins, 1 fibronectin/fibrinogen-binding protein, 1 Chitin-binding protein (located on pLPE10-1 plasmid), 1 ABC superfamily ATP binding cassette transporter, binding protein, 2 cell surface proteins, 1 manganese ABC transporter substrate-binding protein, 1 elongation factor Tu, 1 Molecular chaperone DnaK, 1 molecular chaperone GroEL, 1 co-chaperone GroES, 1 class A sortase and 1 type I glyceraldehyde-3-phosphate dehydrogenase (Table 6). Regarding exopolysaccharides, four genes coding for exopolysaccharide biosynthesis protein were identified (Table 6). For adaptation to different lifestyles, L. pentosus CF2-10N harbored in its genome several genes involved in stress response such as acids and bile. These included three GNAT family acetyltransferases, two Na+/H+ antiporter NhaC, 1 phosphoglycerate mutase, nine elongation factors (factor G, factor GreA, factor 4, factor P, factor Ts and factor Tu) and 1 phosphoglycerate kinase (Table 6).

On the other hand, several genes were identified coding for enzymes involved in probiotic functions such as two genes coding for tannase (exclusive to this strain), 1 alpha-amilase, 1 amylolpullulanase, 3 beta-galactosidases, 5 aminopeptidases, 1 lipase esterase, 4 peptidases, 2 alpha/beta hydrolases, 1 phenolic acid decarboxylase, 1 carboxylesterase, 1 alpha-acetolactate decarboxylase, and 1 multicopper oxidase (Table 6).

With respect to vitamin biosynthesis, we detected genes coding for proteins involved in vitamins B1 or thiamine (10 genes), B2 or riboflavin (8 genes), B5 (3 genes) and B6 (6 genes), folate (7 genes) and vitamin K2 or menaquinone (1 gene) production (Table 6). In this regard, vitamin production ability of L. pentosus CF2-10N was validated in vitro.

Discussion

Aloreña table olives, naturally fermented traditional green olives with a denomination of protection (DOP), are considered as potential source of probiotic L. pentosus strains with high genetic...
| Gene ID | Gene | Position | Strand | Gene length | Protein description | COG ID (COG class) | COG class (COG description) | Similarity to transposase in Lactiplantibacillus pentosus CF2-10N genome |
|---------|------|----------|--------|-------------|---------------------|------------------|-----------------------------|--------------------------------|
| gene_86 | gene_86 | 89,400–90,662 | + | 1,263 | IS3 family transposase | – | – | 99.3% L. pentosus |
| gene_204 | LPENT_00003 | 219,476–220,444 | – | 969 | MULTISPECIES: IS30 family transposase | – | – | 100% Lactobacillaceae |
| gene_638 | gene_638 | 700,336–701,475 | – | 1,140 | Transposase | COG0675 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 100% L. pentosus |
| gene_639 | LPE_01510 | 701,456–701,908 | – | 453 | Transposase family protein A | COG1943 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 100% L. pentosus |
| gene_700 | FD4_1 | 761,249–762,559 | – | 1,311 | IS380 family transposase | – | – | 100% L. pentosus |
| gene_1236 | gene_1236 | 1,345,052–1,345,423 | + | 372 | IS5 family transposase | COG3293 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 100% P. acidilactici |
| gene_1237 | gene_1237 | 1,345,396–1,345,827 | + | 432 | Putative transposase for insertion sequence element IS6031 | COG3293 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 97.9% L. plantarum |
| gene_2023 | top1 | 2,236,745–2,237,608 | – | 924 | MULTISPECIES: IS30 family transposase | COG0675 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 100% Terrabacteria group |
| gene_2025 | FD4_2 | 2,239,842–2,240,885 | + | 1,044 | MULTISPECIES: IS30 family transposase | COG0675 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 99.7% Lactiplantibacillus |
| gene_2321 | HB47_01150 | 2,551,136–2,552,023 | + | 888 | IS30 family transposase | COG0675 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 100% Lactobacillaceae |
| gene_2680 | FD4_3 | 2,918,782–2,920,092 | – | 1,311 | IS380 family transposase | – | – | 100% L. pentosus |
| gene_2707 | HB47_01150 | 2,948,652–2,949,539 | + | 888 | IS30 family transposase | COG0675 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 100% Lactobacillaceae |
| gene_2843 | HB47_01150 | 3,146,814–3,147,701 | + | 888 | IS30 family transposase | COG0675 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 100% Lactobacillaceae |
| gene_3192 | gene_3292 | 3,516,037–3,516,375 | + | 339 | MULTISPECIES: IS5 family transposase | COG3293 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 100% Bacilli |
| gene_3261 | gene_3261 | 3,595,253–3,596,521 | – | 1,269 | Transposase | COG0675 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 100% L. pentosus |
| gene_3262 | LPE_00194 | 3,596,619–3,597,059 | + | 441 | IS200/IS605 family transposase | COG1943 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 100% L. pentosus |
| gene_3455 | FD4_7 | 24,260–25,474 | – | 1,215 | IS21 family transposase | COG0675 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 100% Lactiplantibacillus |
| gene_3464 | FD4_1 | 36,418–37,728 | + | 1,311 | IS380 family transposase | – | – | 100% L. pentosus |
| gene_3465 | LPE_03103 | 38,516–39,514 | + | 999 | IS30 family transposase | COG0675 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 100% L. pentosus |
| gene_3484 | LSE2_2008 | 53,623–54,552 | – | 930 | MULTISPECIES: IS30 family transposase | COG0675 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 99.7% Lactobacillales |
| gene_3486 | gene_3486 | 55,300–57,009 | + | 1710 | DDE transposase | COG0675 (Transposase and inactivated derivatives) | 1. Replication, recombination and repair | 98.9% L. plantarum |

(Continued)
| Gene ID       | Gene     | Position   | Strand | Gene length | Protein description       | COG ID (COG description)                                                                 | COG class (COG class description)                                                                 | Similarity to transposase in Lactiplantibacillus* |
|--------------|----------|------------|--------|-------------|---------------------------|------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------|
| gene_3490    | FD47_    | 59,298-60,512 | --     | 1,215       | IS21 family transposase    | COG4584 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 100% Lactiplantibacillus                        |
| gene_3492    | LPENT_00003 | 62,337-63,305 | --     | 969         | MULTISPECIES: IS30 family transposase | COG2828 (Transposase and inactivated derivatives, IS30 family)                           | 1. (Replication, recombination and repair)                                                     | 100% Lactobacillaceae                          |
| gene_3505    | gene_3505 | 75,427-75,706 | --     | 1,650       | DDE transposase            | COG3666 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 99.5% L. plantarum                            |
| gene_3506    | gene_3506 | 75,204-75,746 | +      | 543         | Transposase                | COG3666 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 100% L. pentosus                              |
| gene_3507    | gene_3507 | 75,794-76,501 | --     | 708         | Transposase, partial       | COG2963 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 100% L. pentosu                                |
| gene_3509    | gene_3509 | 77,298-77,729 | --     | 432         | Putative transposase for insertion sequence element IS6031 | COG3293 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 97.9% L. plantarum                            |
| gene_3510    | gene_3510 | 77,702-78,073 | --     | 372         | IS5 family transposase     | COG3293 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 100% P. acidilactici                          |
| gene_3519    | FD47_    | 83,726-84,565 | --     | 840         | ISS1, transposase (Orf2), IS3 family transposase | COG2801 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 100% L. pentosu                                |
| gene_3520    | LPENT_00063 | 84,601-85,323 | --     | 723         | Transposase (transposase, IS3 family protein) | COG2963 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 100% L. pentosu                                |
| gene_3530    | LPENT_00003 | 91,009-91,977 | --     | 969         | MULTISPECIES: IS30 family transposase | COG2828 (Transposase and inactivated derivatives, IS30 family)                           | 1. (Replication, recombination and repair)                                                     | 100% Lactobacillaceae                          |
| gene_3557    | FD47_    | 116,556-117,770 | --    | 1,215       | IS21 family transposase    | COG4584 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 100% Lactiplantibacillus                        |
| gene_3565    | gene_3565 | 4,191-4,487  | +      | 297         | Transposase (plasmid)      | --                                                                                         | 100% L. plantarum                              |                                                 |
| gene_3566    | gene_3566 | 4,782-5,018  | +      | 237         | Transposase (plasmid)      | COG3464 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 100% L. plantarum                              |
| gene_3568    | FC27_    | 6,430-6,975  | +      | 566         | Transposase                | COG3328 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 98.9% L. paraplantarum                         |
| gene_3569    | LPENT_00125 | 7,062-7,991  | +      | 930         | Transposase TraISLpl1 (IS30 family) | COG2828 (Transposase and inactivated derivatives, IS30 family)                           | 1. (Replication, recombination and repair)                                                     | 99.7% L. pentosu                                |
| gene_3574    | gene_3574 | 10,980-11,261 | --     | 282         | Transposase IS66 (IS66 family transposase) | --                                                                                         | 100% L. pentosu                                |                                                 |
| gene_3576    | gene_3576 | 11,507-12,520 | --     | 1,014       | IS66 family transposase    | COG3466 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 100% L. pentosu                                |
| gene_3577    | gene_3577 | 12,731-12,940 | --     | 210         | MULTISPECIES: transposase   | COG3466 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 100% L. pentosu                                |
| gene_3579    | gene_3579 | 13,597-14,094 | +      | 498         | Transposase                | COG3293 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 100% L. pentosu                                |
| gene_3603    | gene_3603 | 36,269-36,484 | +      | 216         | MULTISPECIES: transposase   | COG3464 (Transposase and inactivated derivatives)                                         | 1. (Replication, recombination and repair)                                                     | 100% L. pentosu                                |
| gene_3607    | FC99_    | 40,024-40,968 | +      | 945         | IS30 family transposase    | COG2828 (Transposase and inactivated derivatives, IS30 family)                           | 1. (Replication, recombination and repair)                                                     | 100% L. pentosu                                |
## Table 1 (Continued)

| Gene ID       | Gene     | Position | Strand | Gene length | Protein description                      | COG ID (COG description) | COG class (COG class description) | Similarity to transposase in *Lactiplantibacillus* |
|---------------|----------|----------|--------|-------------|------------------------------------------|--------------------------|-----------------------------------|-----------------------------------------------|
| gene_3669    | gene_3669 | 42,455-42,751 | +      | 297         | Transposase (plasmid)                    | –                        | –                                 | 100% L. plantarum                              |
| gene_3610    | gene_3610 | 43,046-43,282 | +      | 237         | Transposase (plasmid)                    | COG3464 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 100% L. plantarum                              |
| gene_3612    | FC27_GL001295 | 44,694-45,239 | +      | 546         | Transposase                              | COG3328 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 98.9% L. paraplanulans                          |
| gene_3613    | LPENT_001225 | 45,526-46,255 | +      | 930         | Transposase                           | COG2826 (Transposase and inactivated derivatives, IS30 family) | L (Replication, recombination and repair) | 99.7% L. pentosus                              |
| gene_3619    | gene_3619 | 49,243-49,524 | −      | 282         | Transposase IS66 (IS66 family)          | –                        | –                                 | 98.9% L. pentosus                              |
| gene_3621    | gene_3621 | 50,051-50,782 | −      | 732         | IS66 family transposase                  | COG2348 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 100% L. pentosus                              |
| gene_3622    | gene_3622 | 50,092-51,201 | −      | 210         | mutispecies multistepicfamily transposase | COG2348 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 100% L. pentosus                              |
| gene_3624    | gene_3624 | 51,857-52,252 | +      | 396         | Transposase                            | COG2309 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 100% L. pentosus                              |
| gene_3641    | gene_3641 | 4,323-4,721   | −      | 399         | Transposase                            | COG2309 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 100% Bacilli                                  |
| gene_3643    | gene_3643 | 4,784-5,122   | −      | 359         | mutispecies iss family transposase      | COG2309 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 99.3% L. pentosus                              |
| gene_3652    | gene_3652 | 12,866-13,795 | +      | 930         | IS60 family transposase                  | COG2826 (Transposase and inactivated derivatives, IS30 family) | L (Replication, recombination and repair) | 100% Lactiplantilaceae                         |
| gene_3653    | NIB07_01110 | 18,988-19,885 | −      | 888         | IS60 family transposase                  | COG2826 (Transposase and inactivated derivatives, IS30 family) | L (Replication, recombination and repair) | 100% Lactiplantilaceae                         |
| gene_3655    | FD47_GL000486 | 20,262-21,576 | +      | 1,215       | IS11 family transposase                  | COG4894 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 100% Lactiplantilaceae                         |
| gene_3657    | gene_3657 | 23,236-23,574 | −      | 359         | mutispecies iss family transposase      | COG2309 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 100% Lactiplantilaceae                         |
| gene_3659    | gene_3659 | 24,120-24,794 | −      | 675         | Transposase                            | COG2415 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 100% Lactiplantilaceae                         |
| gene_3660    | FD05_GL002377 | 27,793-29,125 | +      | 1,341       | IS2 family transposase                  | COG2464 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 100% Lactiplantilaceae                         |
| gene_3675    | FD47_GL002738 | 28,210-30,899 | −      | 840         | IS60 family transposase                  | COG2801 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 100% Lactiplantilaceae                         |
| gene_3673    | LPENT_000036 | 35,125-36,847 | −      | 723         | Transposase                           | COG2826 (Transposase and inactivated derivatives, IS30 family) | L (Replication, recombination and repair) | 100% Lactiplantilaceae                         |
| gene_3680    | gene_3680 | 44,896-45,294 | −      | 599         | Transposase                            | COG2309 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 100% L. pentosus                              |
| gene_3680    | gene_3680 | 47,357-47,605 | −      | 359         | mutispecies iss family transposase      | COG2309 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 100% Lactiplantilaceae                         |
| gene_3689    | gene_3689 | 55,438-56,567 | +      | 930         | IS60 family transposase                  | COG2826 (Transposase and inactivated derivatives, IS30 family) | L (Replication, recombination and repair) | 99.4% L. pentosus                              |
| gene_3700    | NIB07_01110 | 61,570-62,457 | −      | 888         | IS60 family transposase                  | COG2826 (Transposase and inactivated derivatives, IS30 family) | L (Replication, recombination and repair) | 100% Lactiplantilaceae                         |
| gene_3700    | gene_3700 | 62,934-64,148 | +      | 1,215       | IS11 family transposase                  | COG4894 (Transposase and inactivated derivatives) | L (Replication, recombination and repair) | 100% Lactiplantilaceae                         |
| gene_3710    | gene_3710 | 65,413-65,502 | −      | 180         | Transposase                            | –                        | –                                 | 95.2% L. plantarum                              |

* the best hit was indicated.
† sequences of pLPE10-1 plasmid.
‡ sequences of pLPE10-4 plasmid.
§ sequences of pLE10-2 plasmid.

Diversity (Abriouel et al., 2012). Several *L. pentosus* strains isolated from Aloreña table olives throughout the fermentation process were shown to be potential probiotics, with *L. pentosus* MP-10, *L. pentosus* CF1-6 and *L. pentosus* CF2-10N as the best candidates (Pérez...
TABLE 2. Characterization of IS elements found within the genome of *Lactiplantibacillus pentosus* CF2-10N using the ISfinder search tool.

| Sequences producing significant alignments | IS Family | Group | Origin | Score (bits) | E value |
|--------------------------------------------|-----------|-------|--------|--------------|---------|
| ISP1                                       | ISL3      |       | Lactobacillus plantarum | 2,547 | 0.0      |
| ISLd3                                      | IS30      |       | Lactobacillus delbrueckii | 1705 | 0.0      |
| ISLhe30                                    | IS30      |       | Lactobacillus helveticus | 1,635 | 0.0      |
| ISLpl3                                     | IS5       | IS427 | Lactobacillus plantarum | 1,429 | 0.0      |
| ISLxa1                                     | IS30      |       | Lactobacillus sakei | 494 | 6e-136   |
| ISLpl2                                     | IS3       | IS150 | Lactobacillus plantarum | 56.0 | 4e-04    |
| ISLhe65                                    | IS200/IS605 | IS1341 | Lactobacillus helveticus | 54.0 | 0.002    |
| ISP2                                       | IS1182    |       | Lactobacillus plantarum | 52.0 | 0.007    |
| ISMmu1                                     | IS200/IS605 | IS605 | Mitsuokella multacida | 52.0 | 0.007    |
| ISLpl5                                     | IS200/IS605 | IS605 | Lactobacillus johnsonii | 52.0 | 0.007    |
| ISSpm5                                     | IS380     |       | Streptococcus pneumoniae | 50.1 | 0.026    |
| IS1161                                     | IS30      |       | Streptococcus salivarius | 48.1 | 0.10     |
| IS1139                                     | IS30      |       | Streptococcus salivarius | 48.1 | 0.10     |
| IS1086                                     | IS30      |       | Ralstonia eutropha | 48.1 | 0.10     |
| ISRhr6                                     | IS5       | IS1031 | Rhodopirillum rubrum | 46.1 | 0.41     |
| ISAss45                                    | IS3       | IS3   | Arthrobacter arilaitensis | 46.1 | 0.41     |
| ISMam7                                     | IS3       | IS3   | Mycobacterium smegmatis | 46.1 | 0.41     |
| IS670                                      | IS30      |       | Enterococcus faecalis | 46.1 | 0.41     |
| IS1648                                     | IS3       | IS427 | Streptomyces coelicolor | 46.1 | 0.41     |
| ISAe41                                     | IS1595    | ISSoft11 | Actinobacteria bacterium | 44.1 | 1.6   |
| ISBsp5                                     | IS1182    |       | Bacillus sp. | 44.1 | 1.6   |
| ISRam1                                     | IS3       | IS150 | Burkholderia ambifaria | 44.1 | 1.6   |
| ISLrb4                                     | ISLre2    |       | Lactobacillus rhamnosus | 44.1 | 1.6   |
| ISSav4                                     | IS701     |       | Streptomyces avermitilis | 44.1 | 1.6   |
| IS231J                                     | IS4       | IS231 | Bacillus thuringiensis | 44.1 | 1.6   |
| ISCle26                                    | IS110     | IS1111 | Citrobacter freundii | 42.1 | 6.4   |
| ISAb16                                     | IS3       | IS150 | Acinetobacter baumannii | 42.1 | 6.4   |
| ISVal2                                     | IS256     | IS1249 | Veillonella atytica | 42.1 | 6.4   |
| ISDhal15                                   | IS1634    |       | Desulfitobacterium dictionoeliminans | 42.1 | 6.4   |
| ISPan1                                     | IS5       | IS903 | Pantoea ananatis | 42.1 | 6.4   |
| ISPph2                                     | IS630     |       | Pelodictyon phaeoceltatiforme | 42.1 | 6.4   |
| ISShes12                                   | IS1634    |       | Shewanella sp. | 42.1 | 6.4   |
| ISSsec13                                   | IS5       | IS427 | Synchococcus sp. | 42.1 | 6.4   |
| ISEnfa364                                  | IS30      |       | Enterococcus faecalis | 42.1 | 6.4   |
| ISNwi3                                     | IS595     | ISNw1 | Nitrobacter winogradskyi | 42.1 | 6.4   |
| ISMma18                                    | IS1634    |       | Methanosarcina mazei | 42.1 | 6.4   |
| ISCle1                                     | IS607     |       | Campylobacter fetsus | 42.1 | 6.4   |
| ISLpl1                                     | IS30      |       | Lactobacillus plantarum | 42.1 | 6.4   |
| IS987                                      | IS3       | IS51  | Mycobacterium bovis | 42.1 | 6.4   |
| IS986                                      | IS3       | IS51  | Mycobacterium tuberculosis | 42.1 | 6.4   |
| IS6110                                     | IS3       | IS51  | Mycobacterium tuberculosis | 42.1 | 6.4   |
| IS231B                                     | IS4       | IS231 | Bacillus thuringiensis | 42.1 | 6.4   |
| IS231A                                     | IS4       | IS231 | Bacillus thuringiensis | 42.1 | 6.4   |
| IS231K                                     | IS4       | IS231 | Bacillus cereus | 42.1 | 6.4   |
| IS1476                                     | ISL3      |       | Enterococcus faecium | 42.1 | 6.4   |
| Region | Region length | Completeness* | Score | Region position | Localization | Most common phage | GC%  | Total proteins |
|--------|---------------|---------------|-------|----------------|-------------|------------------|------|----------------|
| 1      | 15 kb         | Incomplete    | 30    | 2,227,200–2,242,287 | Chromosome  | PHAGE_Lactob_PLE3_ NC_031125(1) | 41.26 | 11             |
| 2      | 39.9 kb       | Intact        | 150   | 2,260,786–2,300,777 | Chromosome  | PHAGE_Lactob_Sha1_ NC_019489(26) | 41.55 | 54             |
| 3      | 47.7 kb       | Intact        | 150   | 2,808,177–2,855,881 | Chromosome  | PHAGE_Lactob_Sha1_ NC_019489(22) | 41.88 | 68             |
| 4      | 23.4 kb       | Incomplete    | 60    | 39,960–63,408       | pLPE10-1    | PHAGE_Enteroc_ bAA91_ss_ NC_022750(2) | 38.27 | 22             |
| 5      | 13.7 kb       | Questionable  | 80    | 309–14,094          | pLPE10-4    | PHAGE_Staphy_ SPbeta_like_ NC_029119(2) | 34.83 | 26             |
| 6      | 19.4 kb       | Questionable  | 90    | 34,863–54,348       | pLPE10-4    | PHAGE_Staphy_ SPbeta_like_ NC_029119(2) | 40.70 | 38             |
| 7      | 18.8 kb       | Questionable  | 80    | 17,673–36,515       | pLPE10-2    | PHAGE_Escher_500,465_1_ NC_049342(3) | 41.54 | 22             |
| 8      | 6.7 kb        | Incomplete    | 40    | 60,245–66,978       | pLPE10-2    | PHAGE_Escher_500,465_1_ NC_049342(3) | 31.68 | 10             |

* Intact (score > 90), Questionable (score 70 ± 90), Incomplete (score <70).
pLPE10: plasmid of L. pentosus CF2-10.
| Gene ID    | Gene     | Position            | Strand | Gene length (bp) | Protein description                                                                 | Ontology ID                                                                 | Ontology term                                                                 |
|-----------|----------|---------------------|--------|------------------|-------------------------------------------------------------------------------------|------------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| gene_1618 | cas9     | 1,785,693–1,789,037 | +      | 3,345            | Type II CRISPR-associated protein Cas9                                                | GO:0003677, GO:0003723, GO:004519, GO:0046872, GO:0043571, GO:0051607, GO:0090305 | DNA binding, RNA binding, endonuclease activity, metal ion binding, maintenance of CRISPR repeat elements, defense response to virus, nucleic acid phosphodiester bond hydrolysis |
| gene_1619 | cas9     | 1,789,043–1,789,759 | +      | 717              | Type II CRISPR-associated protein Cas9                                                | GO:0003677, GO:0003723, GO:004519, GO:0046872, GO:0043571, GO:0051607, GO:0090305 | DNA binding, RNA binding, endonuclease activity, metal ion binding, maintenance of CRISPR repeat elements, defense response to virus, nucleic acid phosphodiester bond hydrolysis |
| gene_1620 | cas1     | 1,789,953–1,790,858 | +      | 906              | Subtype II CRISPR-associated endonuclease Cas1                                      | GO:0003677, GO:0004519, GO:0046872, GO:0043571, GO:0051607, GO:0090305       | DNA binding, endonuclease activity, metal ion binding, maintenance of CRISPR repeat elements, defense response to virus, nucleic acid phosphodiester bond hydrolysis |
| gene_1621 | cas2     | 1,790,836–1,791,141 | +      | 306              | MULTISPECIES: CRISPR-associated endonuclease Cas2                                   | GO:0004521, GO:0046872, GO:0043571, GO:0051607, GO:0090502                   | Endoribonuclease activity, metal ion binding, maintenance of CRISPR repeat elements, defense response to virus, RNA phosphodiester bond hydrolysis, endonucleolytic |
| gene_1622 | gene_1622 | 1,791,138–1,791,815 | +      | 678              | Type II-A CRISPR-associated protein Cas2                                              | −                                                                              | −                                                                              |
| gene_2923 | cas1     | 3,233,320–3,234,273 | +      | 954              | Subtype I-E CRISPR-associated endonuclease Cas1                                     | GO:0003677, GO:0004519, GO:0046872, GO:0043571, GO:0051607, GO:0090305       | DNA binding, endonuclease activity, metal ion binding, maintenance of CRISPR repeat elements, defense response to virus, nucleic acid phosphodiester bond hydrolysis |
| gene_2924 | FD24_GL002157 | 3,234,270–3,235,169 | +      | 900              | Type I-E CRISPR-associated endonuclease Cas2                                         | GO:0003676                                                                    | Nucleic acid binding                                                              |
| gene_2925 | gene_2925 | 3,236,488–3,239,202 | +      | 2,715            | CRISPR-associated helicase/endonuclease Cas3                                         | −                                                                              | −                                                                              |

(Continued)
Among these strains, *L. pentosus* CF2-10N was selected for a more in-depth analysis in the current study on the basis of its excellent probiotic properties. These include notably good growth capacity and survival under simulated gastrointestinal conditions (acidic pH of 1.5, up to 4% of bile salts and 5 mM of nitrate), good ability to auto-aggregate and co-aggregate with pathogenic bacteria, adherence to intestinal and vaginal cell lines, antimicrobial activity by means of plantaricins and fermentation of prebiotics and lactose (Pérez Montoro et al., 2016). It is also noteworthy that *L. pentosus* CF2-10N was isolated from the same ecological niche as the potential previously described probiotic *L. pentosus* MP-10 (Abriouel et al., 2012), hence, they are exposed to the same ecological conditions and pressure (soil, plant and brine) as well as the same progressive changes throughout the production process. It is thus not surprising that their genetic relatedness is further highlighted by shared genetic, functional and probiotic properties although both strains showed different genomic profiles belonging to different clusters or genomic groups (G1 and G2) as reported by Abriouel et al. (2012). In this sense, both strains harbor a single circular chromosome of similar size of 3,698,214 bp (*L. pentosus* MP-10, GC content of 46.32%) and 3,645,747 bp (*L. pentosus* CF2-10N, GC content of 46.42%) and 4 (L. pentosus CF2-10N, 58–120 kb) to 5 (L. pentosus MP-10, 29–46.5 kb) plasmids (Abriouel et al., 2016). This similarity highlights the effect of the ecosystem (soil, plant and brine) on the genetic diversity of microbial communities present in Aloreña table olives.

A comparison with other bacterial strains from table olives showed similarities in genomic size and GC content. These strains included *L. pentosus* IG1 harboring a circular chromosome of 3,687,424 bp (GC content of 44.9%) and 7 plasmids (2.5–125.9 kb; Maldonado-Barragán et al., 2011), *L. pentosus* strains (IG8, IG9, IG10 and IG11) recovered from biofilms on the skin of green table olives with circular chromosome sizes in the range of 3,787,967 to 3,811,295 bp (GC content of 45.9–45.95%) and 6 to 7 plasmids (Calero-Delgado et al., 2019) and *L. pentosus* O17 isolated from brines of treated table olives (*Cerignola* cv.) with a circular chromosome of 3,850,701 bp (GC content of 45.9%; Zotta et al., 2022). This fact indicated their adaptation to a brine-specific lifestyle notably in relation to genes involved in carbohydrate transport and metabolism (307 CDSs and 279 in *L. pentosus* CF2-10N and MP-10, respectively) and amino acid metabolism.

### Table 5: Characterization of CRISPR arrays predicted in the *Lactiplantibacillus pentosus* CF2-10N genome.

| CRISPR array (CR) | Start position | End position | Array orientation | CRISPR length (bp) | Number of repeats | DR consensus* | Array family |
|------------------|----------------|--------------|------------------|-------------------|-----------------|--------------|-------------|
| CR 1             | 1,791,840      | 1,792,537    | Forward          | 698               | 36              | GTCTTGAAATAGTAGAC | NA          |
|                  |                |              |                  |                   |                 | GGGTTGGAAC    |             |
| CR 2             | 2,982,059      | 2,981,480    | Reverse          | 580               | 28              | CTGTTCGGCTCGCCATAGGGGATCTCC | 1-E        |
| CR 3             | 3,232,173      | 3,231,961    | Reverse          | 213               | 28              | CTATTCGGCTCGCATAGGGGATCTCC | NA         |
| CR 4             | 3,232,764      | 3,232,310    | Reverse          | 455               | 28              | GTGTTCGGCGCTCGGAGGGGATCTCC | NA         |
| CR 5             | 3,235,959      | 3,235,382    | Reverse          | 578               | 28              | GTGTTCGGCGCTCGGAGGGGATCTCC | 1-E        |

*The same DR consensus sequences are indicated.*

### Table 4 (continued)

| Gene ID     | Gene          | Position       | Strand | Gene length (bp) | Protein description                  | Ontology ID             | Ontology term                        |
|-------------|---------------|----------------|--------|------------------|--------------------------------------|-------------------------|---------------------------------------|
| gene_2926*  | gene_2926     | 3,239,207-3,240,958 | +      | 1752             | CRISPR-associated protein             | –                       | –                                     |
| gene_2927*  | gene_2927     | 3,240,948-3,241,559 | +      | 612              | Type I-E CRISPR-associated protein    | –                       | –                                     |
| gene_2928*  | gene_2928     | 3,241,559-3,242,638 | +      | 1,080            | Type I-E CRISPR-associated protein    | –                       | –                                     |
| gene_2929   | FD24_GLM02163 | 3,242,619-3,243,344 | +      | 726              | Type I-E CRISPR-associated protein    | GO:0003723, GO:0043571, GO:0051607 | RNA binding, maintenance of CRISPR repeat elements, defense response to virus |
| gene_2930*  | gene_2930     | 3,243,344-3,244,012 | +      | 669              | Type I-E CRISPR-associated protein    | –                       | –                                     |
| Probiotic property | Gene ID | Gene | Position | Strand | Gene length | Protein description | Ontology term (Ontology ID) | COG class (COG class description) |
|-------------------|--------|------|----------|--------|-------------|--------------------|-----------------------------|----------------------------------|
| Adhesion          | gene_411 | FD24_GL003356 | 445,136–448,294 | –       | 3,159        | Mucus-binding protein | Integral component of membrane (GO:0016021) | COG3846 (Type IV secretion pathway; TolB component(s)) |
|                   | gene_963 | LPE_00710   | 1,054,378–1,060,929 | –       | 6,552        | Mucus-binding protein | Integral component of membrane (GO:0016021) | COG9810 (Periplasmic protein TonB, links inner and outer membranes) |
|                   | gene_3039 | gene_3039 | 3,352,844–3,359,728 | –       | 6,885        | Mucus-binding protein | – | COG5099 (RNA-binding protein of the Puf family, translational repressor) |
|                   | gene_3173 | gene_3173 | 3,497,668–3,499,374 | –       | 1,707        | Fibronectin/ fibrinogen-binding protein | – | COG1293 (Predicted RNA-binding protein homologous to eukaryotic snRNP) |
|                   | gene_3512 | gene_3512 | 78,678–78,812 | +       | 135          | Chitin-binding protein | – | COG3397 (Uncharacterized protein conserved in bacteria) |
|                   | gene_891 | LPE_02200   | 975,971–976,864 | –       | 894          | ABC superfamily ATP binding cassette transporter, binding protein | Metal ion binding, cell adhesion, metal ion transport (GO:0006872, GO:0007155, GO:0030001) | COG4803 (ABC-type metal ion transport system, periplasmic component/surface adhesin) |
|                   | gene_517 | LPE_00567   | 561,619–563,421 | +       | 1,803        | Cell surface protein | Extracellular region, cell wall, integral component of membrane, collagen binding, cell adhesion (GO:0005576, GO:0005618, GO:0016021, GO:0005518, GO:0007155) | COG4932 (Predicted outer membrane protein) |
|                   | gene_840 | FD24_GL00462 | 920,457–922,340 | –       | 1884         | Cell surface protein | Extracellular region, cell wall, collagen binding, cell adhesion (GO:0005576, GO:0005618, GO:0005518, GO:0007155) | COG8903 (ABC-type metal ion transport system, periplasmic component/surface adhesin) |
|                   | gene_2496 | FD24_GL00106 | 2,735,640–2,736,581 | +       | 942          | Manganese ABC transporter substrate-binding protein | Metal ion binding, cell adhesion, metal ion transport (GO:0005576, GO:0005618, GO:0005518, GO:0007155) | COG4803 (ABC-type metal ion transport system, periplasmic component/surface adhesin) |
|                   | gene_158 | tuf      | 162,869–164,056 | –       | 1,188        | Elongation factor Tu | Cytoplasm, translation elongation factor activity, GTPase activity, GTP binding, translational elongation (GO:0005737, GO:0005746, GO:0003924, GO:0006527, GO:0006414) | COG4585 (GTPases, translation elongation factors) |
|                   | gene_74  | dnaK      | 74,096–75,964 | –       | 1,869        | Molecular chaperone Dnak | ATP binding, unfolded protein binding, protein folding (GO:0005524, GO:0001082, GO:0006457) | COG4843 (Molecular chaperone) |
|                   | gene_2181 | groL      | 2,382,568–2,384,193 | +       | 1,626        | MULTISPECIES: molecular chaperone GroEL | Cytoplasm, ATP binding, unfolded protein binding, protein refolding (GO:0005737, GO:0005524, GO:0005108, GO:0004206) | COG4589 [Chaperonin GroEL (HSP60 family)] |
|                   | gene_2180 | groS      | 2,382,228–2,382,512 | +       | 285          | MULTISPECIES: co-chaperone GroES | Cytoplasm, ATP binding, protein folding (GO:0005737, GO:0005524, GO:00046457) | COG0234 [Co-chaperonin GroES (HSP10)] |
|                   | gene_1964 | N692_13295 | 2,164,82–2,165,546 | +       | 705          | MULTISPECIES: class A surfase | Integral component of membrane (GO:0010201) | COG3784 [Surfase (surface protein transport/pilus)] |
|                   | gene_2239 | LPENT_01088 | 2,455,749–2,456,771 | +       | 1,023        | MULTISPECIES: type I glycerolaldehyde-3-phosphate dehydrogenase (NAD+) (phosphorylating) activity, NADP binding, NAD binding, glucose metabolic process, oxidation–reduction process (GO:0004365, GO:0050661, GO:0031287, GO:0006006, GO:0005114) | COG0507 [Glycerolaldehyde-3-phosphate dehydrogenase/cytochrome 4-phosphate dehydrogenase] |

(Continued)
### TABLE 6 (Continued)

| Probiotic property | Gene ID | Gene | Position | Strand | Gene length | Protein description | Ontology term (Ontology ID) | COG class (COG class description) |
|--------------------|---------|------|----------|--------|-------------|---------------------|-----------------------------|----------------------------------|
|                     |         |      |          |        |             |                     |                             |                                  |
| **Exopolysaccharides** |         |      |          |        |             |                     |                             |                                  |
| gene_346            | LPE_00940 | 151,375–152,091 | −   | 717 | Exopolysaccharide biosynthesis protein | Extracellular polysaccharide biosynthetic process (GO:0045226) | COG0489 (ATPases involved in chromosome partitioning) |
| gene_2631           | LPE_02641 | 2,677,199–2,687,827 | +   | 720 | Exopolysaccharide biosynthesis protein | Extracellular polysaccharide biosynthetic process (GO:0045226) | COG0489 (ATPases involved in chromosome partitioning) |
| gene_2651           | LPE_01025 | 2,677,199–2,687,827 | +   | 720 | Exopolysaccharide biosynthesis protein | Extracellular polysaccharide biosynthetic process (GO:0045226) | COG0489 (ATPases involved in chromosome partitioning) |
| gene_2676           | LPE_00805 | 2,613,577–2,614,353 | +   | 777 | Exopolysaccharide biosynthesis protein | Transmembrane transport (GO:0016021, GO:0015297, GO:0055085) | COG1757 (Na+/H+ antiporter) |

**Exopolysaccharides**

| Gene | Position | Strand | Gene length | Protein description | Ontology term (Ontology ID) | COG class (COG class description) |
|------|----------|--------|-------------|---------------------|-----------------------------|----------------------------------|
| gene_74   | dnaK    | 74,096–75,964 | −   | 1,869 | Molecular chaperone DnaK | ATP binding, unfolded protein binding, protein folding (GO:0016021, GO:0005506, GO:0005524) | COG3953 (Molecular chaperones) |
| gene_682  | pyrD    | 664,092–665,009 | −   | 918  | Dihydroorotate dehydrogenase B catalytic subunit | Cytosolic, dihydroorotate dehydrogenase activity, 'de novo' pyrimidine nucleobase biosynthetic process, 'de novo' UMP biosynthetic process, oxidation–reduction process (GO:0005737, GO:0006457, GO:0004152, GO:0055114) | COG1677 (Dihydroorotate dehydrogenase) |
| gene_688  | LPE_01835 | 1,209,106–1,209,562 | +   | 458  | N-acetyltransferase activity | N-acetyltransferase activity (GO:0008080) | COG4552 (Predicted acetyltransferases involved in intracellular survival and related acetyltransferases) |
| gene_670  | LPE_01051 | 1,091,017–1,091,502 | −   | 489  | N-acetyltransferase activity | N-acetyltransferase activity (GO:0008080) | COG4552 (Predicted acetyltransferases involved in intracellular survival and related acetyltransferases) |
| gene_172  | FD24_G2001267 | 1,204,765–1,206,187 | −   | 1,422 | Na+/H+ antiporter NhaC | Integral component of membrane, antiporter activity, transmembrane transport (GO:0016021, GO:0015297, GO:0055085) | COG1757 (Na+/H+ antiporter) |
| gene_1684 | LPE_02288 | 1,065,409–1,066,838 | −   | 1,401 | Na+/H+ antiporter NhaC | Integral component of membrane, antiporter activity, transmembrane transport (GO:0016021, GO:0015297, GO:0055085) | COG1757 (Na+/H+ antiporter) |
| gene_2117 | gpmA    | 2,321,252–2,321,899 | +   | 678  | Phosphoglycerate mutase | 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase activity, gluconeogenesis, glycolytic process (GO:0006538, GO:0006593, GO:0006594, GO:0006595) | COG0586 (Phosphoglycerate mutase) |
| gene_2146 | groEL   | 2,302,568–2,304,933 | +   | 1,026 | MULTISPECIES: molecular chaperone GroEL | Cytosolic ATP binding, unfolded protein binding, protein refolding (GO:0006537, GO:0006524, GO:0015297, GO:0005506, GO:0005524) | COG0679 (Chaperonin GroEL (HSP60 family) |
| gene_2223 | luxS    | 2,638,407–2,639,283 | +   | 477  | MULTISPECIES: S-ribosylhomocysteine lyase | S-ribosylhomocysteine lyase activity, quorum sensing (GO:0005536, GO:0043768, GO:0009372) | COG1854 (LuxS protein involved in autoinducer AI2 synthesis) |
| gene_2434 | fusA    | 2,682,265–2,683,475 | +   | 2,097 | MULTISPECIES: elongation factor G | Cytosolic, translation elongation factor activity, GTPase activity, GTP binding, translational elongation (GO:0006537, GO:0006546, GO:003926, GO:0005523, GO:0004414) | COG0480 ([Translation elongation factors (GTPases)]) |
| Probiotic property | Gene ID | Gene  | Position          | Strand | Gene length | Protein description | Ontology term (Ontology ID)                                                                 | COG class (COG class description) |
|--------------------|---------|-------|-------------------|--------|-------------|--------------------|---------------------------------------------------------------------------------------------|----------------------------------|
| gene_2964          | greA    |       | 3,282,554–3,283,014 | +      | 483         | MULTISPECIES: transcription elongation factor GreA DNA binding, translation elongation factor activity, RNA polymerase binding, transcription, DNA-templated, translational elongation, regulation of DNA-templated transcription, elongation (GO:0003677, GO:0003746, GO:0037003, GO:0006351, GO:0036416, GO:0032784) | COG0762 (Transcription elongation factor) |
| gene_1712          | greA    |       | 1,893,369–1,893,839 | −      | 471         | Transcription elongation factor GreA DNA binding, translation elongation factor activity, RNA polymerase binding, transcription, DNA-templated, translational elongation, regulation of DNA-templated transcription, elongation (GO:0003677, GO:0003746, GO:0037003, GO:0006351, GO:0036416, GO:0032784) | COG0762 (Transcription elongation factor) |
| gene_2240          | pgk     |       | 2,656,860–2,658,091 | +      | 1,203       | MULTISPECIES: phosphoglycerate kinase DNA binding, translation elongation factor activity, RNA polymerase binding, transcription, DNA-templated, translational elongation, regulation of DNA-templated transcription, elongation (GO:0003677, GO:0003746, GO:0070063, GO:0006351, GO:0006414, GO:0032784) | COG0126 (3-phosphoglycerate kinase) |
| gene_66            | lepA    |       | 64,623–66,458      | −      | 1,836       | Elongation factor 4 Plasma membrane, translation elongation factor activity, GTPase activity, GTP binding, ribosome binding, translational elongation, positive regulation of translation (GO:0005886, GO:0035376, GO:0039326, GO:0005523, GO:0043022, GO:0036414, GO:0045727) | COG0481 (Membrane GTPase LepA) |
| gene_1072          | lepA    |       | 1,171,570–1,173,357 | +      | 1,788       | Elongation factor 4 Plasma membrane, translation elongation factor activity, GTPase activity, GTP binding, ribosome binding, translational elongation, positive regulation of translation (GO:0005886, GO:0035376, GO:0039326, GO:0005523, GO:0043022, GO:0036414, GO:0045727) | COG0481 (Membrane GTPase LepA) |
| gene_1569          | FD24_GL002972 |     | 1,732,555–1,734,524 | +      | 1,992       | Elongation factor G Translation elongation factor activity, GTPase activity, GTP binding, translational elongation (GO:0035376, GO:0039326, GO:0005523, GO:0043022, GO:0036414, GO:0045727) | COG0480 (Translation elongation factors (GTPases)) |
| gene_2996          | efp     |       | 3,308,149–3,308,706 | +      | 558         | MULTISPECIES: elongation factor P DNA binding, translation elongation factor activity, RNA polymerase binding, transcription, DNA-templated, translational elongation, regulation of DNA-templated transcription, elongation (GO:0003677, GO:0003746, GO:0036416, GO:0032784) | COG0351 \Translation elongation factor P (EF-P)\ |
| gene_158           | FD24_GL003074 |     | 1,841,220–1,842,542 | +      | 1,323       | Elongation factor Ts DNA binding, translation elongation factor activity, RNA polymerase binding, transcription, DNA-templated, translational elongation, regulation of DNA-templated transcription, elongation (GO:0003677, GO:0003746, GO:0036416, GO:0032784) | COG0564 (Translation elongation factor Ts) |
| Enzymes            | gene_11 | gene_11 | 7,006–9,687       | −      | 1884        | Tannase DNA binding, translation elongation factor activity, RNA polymerase binding, transcription, DNA-templated, translational elongation, regulation of DNA-templated transcription, elongation (GO:0003677, GO:0003746, GO:0036416, GO:0032784) | COG0550 (Tannase, translation elongation factors) |
| gene_3283          | gene_3283 |     | 3,633,861–3,635,744 | −      | 1884        | Tannase DNA binding, translation elongation factor activity, RNA polymerase binding, transcription, DNA-templated, translational elongation, regulation of DNA-templated transcription, elongation (GO:0003677, GO:0003746, GO:0036416, GO:0032784) | COG0550 (Tannase, translation elongation factors) |
| gene_1672          | FD24_GL003074 |     | 1,841,220–1,842,542 | +      | 1,323       | Alpha-amylase DNA binding, translation elongation factor activity, RNA polymerase binding, transcription, DNA-templated, translational elongation, regulation of DNA-templated transcription, elongation (GO:0003677, GO:0003746, GO:0036416, GO:0032784) | COG0550 (Tannase, translation elongation factors) |
| Gene ID    | Gene     | Position          | Strand | Gene length | Protein description                                                                 | Ontology term (Ontology ID)                                                                 | COG class (COG class description) |
|------------|----------|-------------------|--------|-------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|----------------------------------|
| gene_1516  | LPE_01041| 1,670,552–1,681,369| +      | 1818        | Amylopullulanase                                                                     | Alpha-amylase activity, carbohydrate metabolic process (GO:0004556, GO:0005975)               | COG0506 (Glycosidase)             |
| gene_1271  | FDQG_G250181 | 1,379,908–1,381,859  | +      | 1934        | Beta-galactosidase                                                                  | Beta-galactosidase complex, beta-galactosidase activity, metal ion binding, galactose metabolic process (GO:0003941, GO:0004565, GO:0406872, GO:0005975) | COG1874 (Beta-galactosidase)     |
| gene_1284  | FDQG_G250180 | 1,394,185–1,396,065  | +      | 1881        | Beta-galactosidase                                                                  | Beta-galactosidase complex, beta-galactosidase activity, carbohydrate metabolic process (GO:0003941, GO:0004565, GO:0005975) | COG3250 (Beta-galactosidase/beta-glucuronidase) |
| gene_1285  | FDQG_G250187 | 1,396,049–1,397,008  | +      | 960         | Beta-galactosidase                                                                  | Beta-galactosidase complex, beta-galactosidase activity, carbohydrate metabolic process (GO:0003941, GO:0004565, GO:0005975) | COG3250 (Beta-galactosidase/beta-glucuronidase) |
| gene_1422  | FDQG_G250186 | 1,558,889–1,561,432  | +      | 2,544       | Hypothetical protein                                                                | Beta-galactosidase complex, beta-galactosidase activity, carbohydrate metabolic process (GO:0003941, GO:0004565, GO:0005975) | COG1874 (Beta-galactosidase)     |
| gene_1988  | map      | 2,194,635–2,195,540 | −      | 906         | MULTISPECIES: prolyl aminopeptidase                                                 | Aminopeptidase activity, proteolysis (GO:0004177, GO:0006508)                                | COG0596 ([Predicted hydrodrolases or acyltransferases](alpha/beta-hydrolase superfamily)) |
| gene_1749  | map      | 1,026,072–1,026,863 | −      | 792         | MULTISPECIES: type I methionyl aminopeptidase                                       | Metal ion binding, metalloaminopeptidase activity, proteolysis, protein initiator methionine removal (GO:0406872, GO:0006508, GO:0406872, GO:0006508) | COG0024 (Metalloaminopeptidase)    |
| gene_2295  | FDQG_G250273 | 2,525,861–2,526,749  | −      | 908         | Prolyl aminopeptidase                                                               | Aminopeptidase activity, proteolysis (GO:0004177, GO:0006508)                                | COG0596 ([Predicted hydrodrolases or acyltransferases](alpha/beta-hydrolase superfamily)) |
| gene_3445  | LPE_01265 | 1,634,935–1,636,251 | −      | 317         | Aminopeptidase                                                                      | Aminopeptidase activity, cysteine-type endopeptidase activity, proteolysis (GO:04064177, GO:0006508) | COG0579 (Aminopeptidase C)       |
| gene_2120  | LDVNT_01205 | 2,525,277–2,526,898  | −      | 1,352       | Aminopeptidase                                                                      | Aminopeptidase activity, cysteine-type endopeptidase activity, proteolysis (GO:04064177, GO:0006508) | COG0579 (Aminopeptidase C)       |
| gene_2364  | FDQG_G250247 | 2,600,985–2,603,517  | +      | 2,535       | Peptidase                                                                           | Aminopeptidase activity, metallopeptidase activity, zinc ion binding, metallopeptidase activity (GO:04064417, GO:0006508) | COG0508 (Aminopeptidase N)       |
| gene_3265  | pepQ      | 3,307,014–3,308,878  | +      | 1,239       | Peptidase M24 family protein                                                        | Hydrolase activity (GO:017478)                                                             | COG0006 (Xaa-Pro aminopeptidase)   |
| gene_3297  | LPE_00653 | 3,636,220–3,639,118 | −      | 891         | Alpha/beta hydrolase                                                                | (GO:0004177, GO:0006508)                                                                     | COG1506 (Dipeptidyl aminopeptidase/acylaminopeptidases) |
| Probiotic property | Gene ID       | Gene    | Position       | Strand | Gene length | Protein description                                                                 | Ontology term (Ontology ID)                                                                 | COG class (COG class description)                                                                 |
|-------------------|---------------|---------|----------------|--------|-------------|-------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------|
| gene_15           | LPE_00163     |         | 12,171–13,061  | −      | 891         | Alpha/beta hydrolase                                                                | serine-type peptidase activity, proteolysis                                                   | COG1506 (Dipeptidyl aminopeptidase/ace/aminopeptidase)                                                                    |
| gene_1479         | FD24_GEO000907|         | 1,630,300–1,630,836| +      | 537         | Phenolic acid decarboxylase                                                          | carboxy-lyase activity (GO:0016831)                                                          | COG2259 (Phenolic acid decarboxylases)                                                                                     |
| omc_2346          | LPE_001287    |         | 2,463,696–2,464,344| +      | 747         | Carbamoylase                                                                        | Carbamoyl group translocase activity, carbamoyl-removal process (GO:0005288)                  | COG1557 (Enzyme family)                                                                                                    |
| gene_77           | FD24_GEO001463|         | 78,101–78,911   | −      | 711         | Alpha-acetolactate decarboxylase                                                     | Amino acid decarboxylase activity, amino acid biosynthetic process (GO:0005285, GO:0013611) | COG1557 (Alpha-acetolactate decarboxylase)                                                                                   |
| gene_968          | FD24_GEO000152|         | 883,614–884,444 | −      | 831         | Lipase                                                               | Hydrolyase activity, metabolism (GO:0003778, GO:0004352)                                        | COG0857 (Enzyme family)                                                                                                    |
| gene_1852         | LPE_000668    |         | 2,040,108–2,041,613| +      | 1,506       | MULTISPECIES: multicopper oxidase                                                   | Copper ion binding, oxidoreductase activity, cell division, oxidation-reduction process (GO:0005587, GO:0116491, GO:0106140, GO:0351141) | COG2152 (Putative multicopper oxidase)                                                                                     |
|                   |               |         |                |        |             | Vitamin                                                                            |                                                                                                  |                                                                                                                          |
| Vitamins          | Folate        | gene_335| FD24_GEO000368 | −      | 1,326       | Betalactyl-phosphatase synthase/dihydrofolate synthase                              | Tetrahydrofolate synthase activity, ATP binding, tetrahydrofolate biosynthetic process (GO:0004326, GO:0005524, GO:0004881) | COG2065 (Betalactyl-phosphatase synthase)                                                                                   |
| gene_1,140        | LPE_01427     |         | 1,241,801–1,243,829| −      | 1,249       | Dihydropyruvate synthase                                                            | Dihydropyruvate synthase activity, folic acid-containing compound biosynthetic process (GO:0004736, GO:0008938) | COG2084 (Dihydropyruvate synthase and related enzymes)                                                                        |
| gene_1,145        | LPENT_02001    |         | 1,246,075–1,246,443| −      | 369         | Dihydronopterin aldolase                                                             | Dihydronopterin aldolase activity, tetrahydrofolate biosynthetic process, folic acid biosynthetic process (GO:0004530, GO:0004634, GO:0004656) | COG1739 (Dihydronopterin aldolase)                                                                                           |
| gene_3153         | Fhu           | 3,480,636–3,482,289| +      | 1,656       | Formate−tetrahydrofolate ligase                                                      | Formate−tetrahydrofolate ligase activity, folic acid-containing compound biosynthetic process, tetrahydrofolate interconversion (GO:0008282, GO:0005524, GO:0008938, GO:0003999) | COG2778 (Formate−tetrahydrofolate synthase)                                                                                   |
| gene_3143         | JDE           | 1,245,011–1,245,580| −      | 570         | MULTISPECIES: GTP cytochrome oxidase I activity, GTP binding, iron ion binding, iron-carrying metalloprotein, 7-α-hydroxysterol synthase 3′-phosphoglycerate biosynthetic process, tetrahydrofolate biosynthetic process (GO:0005737, GO:0003954, GO:0005525, GO:0008870, GO:0006708, GO:0005594, GO:0046596) | COG0301 (GTP-cytochrome oxidase)                                                                                         |
| gene_3144         | LPE_01431     |         | 1,245,573–1,246,085| −      | 513         | 2-amino-6-hydroxy-6-hydroxymethylpyrophosphatase                                   | 2-amino-6-hydroxy-6-hydroxymethylpyrophosphatase activity, kinase activity, folic acid-containing compound biosynthetic process, phosphatase (GO:0005488, GO:0043410, GO:0005978, GO:0005610) | COG0891 (7,8-hydroxymethyl-7,8-dihydro-5-oxoporphyrin-5-ylphosphate synthase)                                               |
### Table 6 (Continued)

| Probiotic property | Gene ID | Gene | Position | Strand | Gene length | Protein description | Ontology term (Ontology ID) | COG class (COG class description) |
|-------------------|---------|------|----------|--------|-------------|---------------------|-----------------------------|----------------------------------|
| **Riboflavin**    | gene_2999 | folD | 3,108,775–3,108,835 | +      | 861         | Bifunctional protein folD | Methylenetetrahydrofolate cyclohydrolase activity, methylene-tetrahydrofolate dehydrogenase (NADP+) activity, histidine biosynthetic process, purine nucleotide biosynthetic process, methionine biosynthetic process, folic acid-containing compound biosynthetic process, tetrahydrofolate interconversion, oxidation–reduction process (GO:0004477, GO:004488, GO:0000105, GO:0004488, GO:004486, GO:0035996, GO:0035114) | COG0198 (5,10-methylene-tetrahydrofolate dehydrogenase/Methenyl tetrahydrofolate cyclohydrolase) |
|                   | gene_2293 | LPE_03224 | 2,522,708–2,522,836 | +      | 849         | Bifunctional protein riboflavin kinase | FADN adenylyltransferase activity, kinase activity, riboflavin biosynthetic process, phosphorylation (GO:0003919, GO:0003919, GO:0003919, GO:0003919) | COG0016 (FAD synthase) |
|                   | gene_78  | FDS26_GL001464 | 78,815–79,833 | –      | 999         | Bifunctional riboflavin kinase | FADN adenylyltransferase activity, ATP binding, riboflavin kinase biosynthetic process, phosphorylation process, COG0196 (Riboflavin synthase catalytic subunit) | COG0198 (FAD synthase) |
|                   | gene_2838 | FDS26_GL002070 | 3,140,038–3,140,862 | +      | 1,212       | Bifunctional 3,4-dihydroxy-2-butanone-4-phosphate synthase/GTP cyclohydrolase II | GTP cyclohydrolase II activity, GTP binding, 3,4-dihydroxy-2-butanone-4-phosphate synthase activity, metal ion binding, riboflavin biosynthetic process (GO:0003919, GO:0003919, GO:0003919, GO:0003919) | COG0007 (GTP cyclohydrolase II) |
|                   | gene_2839 | ribH | 3,148,962–3,149,429 | +      | 468         | 6,7-dimethyl-8-ribityllumazine synthase | Riboflavin synthase complex, 6,7-dimethyl-8-ribityllumazine synthase activity, translocase activity, riboflavin biosynthetic process (GO:0003919, GO:0003919, GO:0003919) | COG0554 (Riboflavin synthase beta-chain) |
|                   | gene_2836 | LPE_03375 | 3,138,079–3,139,046 | +      | 1,068       | Riboflavin biosynthesis protein RibA | Zinc ion binding, 5-amino-6-(5-phosphoribosylamino) uracil deiodinase activity, dianion dehydrodihydrofolate reductase activity, riboflavin biosynthetic process, oxidation–reduction process (GO:0003919, GO:0003919, GO:0003919) | COG1985 (Pyrimidine reductase, riboflavin biosynthesis) |
|                   | gene_3254 | gene_3294 | 3,584,670–3,585,035 | –      | 381         | Riboflavin biosynthesis protein RibT | Riboflavin biosynthesis protein RibT | – |
|                   | gene_2837 | LPE_03376 | 3,139,147–3,139,740 | +      | 605         | Riboflavin synthase | Oxidoreductase activity, oxidation–reduction process (GO:0014491, GO:0014491) | COG1037 (Riboflavin synthase alpha chain) |
|                   | gene_728  | LPENT_02502 | 795,849–796,349 | +      | 540         | Dihydrofolate reductase | Integral component of membrane, 5-amino-6-(5-phosphoribosylamino) uracil deiodinase activity, riboflavin biosynthetic process, oxidation–reduction process (GO:0003919, GO:0003919) | COG1262 (Dihydrofolate reductase) |

(Continued)
| Probiotic property | Gene ID | Gene | Position | Strand | Gene length | Protein description | Ontology term (Ontology ID) | COG class (COG class description) |
|-------------------|---------|------|----------|--------|--------------|-----------------------|-----------------------------|----------------------------------|
| Thiamine          | gene_1604 | thiE  | 1,772,972–1,773,028 | +      | 657          | Thiamine phosphate synthase | Magnesium ion binding, thiamine-phosphate diphosphorylase activity, thiamine biosynthetic process, thiamine diposphate biosynthetic process (GO:00000287, GO:0004789, GO:0000228, GO:0009229) | COG0352 (Thiamine monophosphate synthase) |
|                   | LPE_00578 | thiE  | 575,356–577,104  | –      | 1749         | 1-deoxy-D-xylulose-5-phosphate synthase | 1-deoxy-D-xylulose-5-phosphate synthase activity, metal ion binding, thiamine biosynthetic process, terpenoid biosynthetic process, 1-deoxy-D-xylulose 5-phosphate biosynthetic process (GO:0046661, GO:0046672, GO:0049220, GO:0046114, GO:0052865) | COG1154 (Deoxyxylulose-5-phosphate synthase) |
|                   | gene_1604 | thiM  | 1,770,746–1,771,540 | +      | 795          | Hydroxyethylthiazole kinase | Magnesium ion binding, hydroxyethylthiazole kinase activity, ATP binding, thiamine biosynthetic process, thiamine diposphate biosynthetic process, phosphorylation (GO:0000287, GO:0004417, GO:0005524, GO:0009228, GO:0009229, GO:0016310) | COG2145 (Hydroxyethylthiazole kinase, sugar kinase family) |
|                   | gene_2902 | FD24_GL002133 | 3,205,215–3,206,249 | –      | 1,035        | Methylsterin biosynthetic protein BioI | Small protein activating enzyme activity (GO:0008641) | COG0476 (Dinucleotide-utilizing enzymes involved in molybdopterin and thiamine biosynthesis family) |
|                   | gene_1605 | LPE_003369 | 1,771,558–1,772,382 | +      | 825          | MULTISPEECES: | ATP binding, phosphomethylpyrimidine kinase activity, thiamine biosynthetic process, phosphorylation (GO:0000287, GO:0004417, GO:0046872, GO:0049220, GO:0046114, GO:0052865) | COG1564 (Thiamine pyrophosphokinase) |
|                   | LPE_00414 | 3,352,290–3,352,498 | –      | 657          | Thiamine pyrophosphokinase | Thiamine diprophosphokinase activity, ATP binding, thiamine binding, thiamine metabolic process, thiamine diphostate biosynthetic process (GO:0008646, GO:0005224, GO:0046775, GO:0046872, GO:0049220, GO:0046114, GO:0052865) | COG1564 (Thiamine pyrophosphokinase) |
|                   | gene_538 | thiF  | 369,546–370,763  | –      | 1,218        | RNA sulfurtransferase | RNA sulfurtransferase activity, ATP binding, thiamine biosynthetic process, thiamine biosynthetic process, thiamine diphostate biosynthetic process, RNA flux-modification (GO:0003737, GO:0000049, GO:0048040, GO:0000524, GO:0046175, GO:0049220, GO:0046114, GO:0052865) | COG1564 (Thiamine pyrophosphokinase) |
|                   | gene_821 | FD24_GL000641 | 897,794–898,705  | +      | 1,062        | FAD protein FMN transferase | Transferase activity, metal ion binding, protein flavinylation (GO:0016740, GO:0046872, GO:0049220, GO:0046114, GO:0052865) | COG1477 (Membrane-associated lipoprotein involved in thiamine biosynthesis) |
|                   | LPE_02557 | 1,421,909–1,422,865 | –      | 957          | FAD protein FMN transferase | Transferase activity, metal ion binding, protein flavinylation (GO:0016740, GO:0046872, GO:0049220, GO:0046114, GO:0052865) | COG1477 (Membrane-associated lipoprotein involved in thiamine biosynthesis) |
|                   | gene_2469 | FD24_GL000340 | 2,700,017–2,710,120 | +      | 1,113        | FAD protein FMN transferase | Transferase activity, metal ion binding, protein flavinylation (GO:0016740, GO:0046872, GO:0049220, GO:0046114, GO:0052865) | COG1477 (Membrane-associated lipoprotein involved in thiamine biosynthesis) |
| Probiotic property | Gene ID | Gene | Position       | Strand | Gene length | Protein description                                      | Ontology term (Ontology ID)                                                                 | COG class (COG class description) |
|-------------------|---------|------|----------------|--------|-------------|--------------------------------------------------------|-------------------------------------------------------------------------------------------|-----------------------------------|
| Vitamin K2        | gene_1240 | menG | 1,346,872–1,347,585 | +      | 714         | Bifunctional demethylmenaquinone methyltranserase/2-methyl-6-polypropyl-1,4-benzoquinol methylase | Methylenetetrahydrofolate reductase activity, menaquinone biosynthetic process, methylation (GO:0008168, GO:000234, GO:0003229) | COG2226 (Methylase involved in ubiquinone/menaquinone biosynthesis) |
| Vitamin B3        | gene_452 | koAD1 | 486,520–487,494 | –       | 975         | 2-dehydrobutyrate-2-reductase | Pyridoxal phosphate biosynthetic process, oxidation-reduction process (GO:0004733, GO:0010181, GO:0044823, GO:00353114) | COG1895 (Ketopantoate reductase) |
|                   | gene_683 | menV  | 750,594–751,323  | +      | 1,020       | 2-dehydrobutyrate-2-reductase | Pyridoxal phosphate biosynthetic process, oxidation-reduction process (GO:0004733, GO:0010181, GO:0044823, GO:00353114) | COG1895 (Ketopantoate reductase) |
|                   | gene_1840 | LPE_00879 | 2,026,749–2,027,732 | +      | 984         | 2-dehydrobutyrate-2-reductase | Pyridoxal phosphate biosynthetic process, oxidation-reduction process (GO:0004733, GO:0010181, GO:0044823, GO:00353114) | COG1895 (Ketopantoate reductase) |
| Vitamin B6        | gene_1521 | FD24_GL000863 | 1,687,267–1,687,710 | –       | 444         | MULTISPECIES: pyridoxamine 5′-phosphate oxidase | Pyridoxamine-phosphate oxidase activity, pyridoxal phosphate biosynthetic process, oxidation-reduction process (GO:0004733, GO:0010181, GO:0044823, GO:00353114) | - |
|                   | gene_653 | FD24_GL002335 | 711,251–712,435  | –       | 1,185       | Pyridoxal phosphate-dependent aminotransferase | Transaminase activity, pyridoxal phosphate biosynthetic process, oxidation-reduction process (GO:0004843, GO:0035170, GO:0035558) | COG1696 (Bifunctional PLP-dependent enzyme with beta-cystathionase and maltose regulon repressor activities) |
|                   | gene_796 | LPE_03241 | 846,881–847,853  | –       | 1,173       | Pyridoxal phosphate-dependent aminotransferase | Transaminase activity, pyridoxal phosphate biosynthetic process, oxidation-reduction process (GO:0004843, GO:0035170, GO:0035558) | COG1696 (Bifunctional PLP-dependent enzyme with beta-cystathionase and maltose regulon repressor activities) |
|                   | gene_1841 | LPE_00878 | 2,027,735–2,028,907 | +      | 1,173       | Pyridoxal phosphate-dependent aminotransferase | Transaminase activity, pyridoxal phosphate biosynthetic process, oxidation-reduction process (GO:0004843, GO:0035170, GO:0035558) | COG1696 (Bifunctional PLP-dependent enzyme with beta-cystathionase and maltose regulon repressor activities) |
|                   | gene_3128 | LPE_05325 | 3,447,961–3,448,198 | +      | 1,173       | Pyridoxal phosphate-dependent aminotransferase | L-lysine:2-oxoglutarate aminotransferase activity, pyridoxal phosphate biosynthetic process (GO:0040489, GO:0035170, GO:0035558) | COG1696 (Bifunctional PLP-dependent enzyme with beta-cystathionase and maltose regulon repressor activities) |
|                   | gene_2303 | LPE_08235 | 2,538,192–2,539,810 | +      | 894         | Pyridoxine kinase | ATP binding, pyridoxal kinase activity, pyridoxal 5′-phosphate salvaging phosphorylation (GO:0005520, GO:0004978, GO:0004441, GO:00353114) | COG2240 (Pyridoxine/tryptophan/tryptophanase kinase) |

*: the best hit was indicated.
$: sequences of pLPE10-1 plasmid.
$: sequences of pLPE10-4 plasmid.
$: sequences of pLPE10-2 plasmid.
and the great diversity of transposases and IS elements identified by \textit{in silico} analysis of the \textit{L. pentosus} CF2-10N genome indicated a frequent genetic diversification within the \textit{L. pentosus} CF2-10N genome, which is notably higher than in other lactobacilli such as \textit{L. plantarum} WCFS1 (36 genes), \textit{L. pentosus} KCA1 (25 genes), \textit{L. pentosus} DSM 20314 (14 genes) or \textit{L. pentosus} IG1 (5 genes; Abriouel et al., 2017). Interestingly, \textit{L. pentosus} CF2-10N showed an even higher genetic diversification in comparison to \textit{L. pentosus} MP-10 (29 genes), even though both strains are isolated from the same ecological niche (Abriouel et al., 2017). Furthermore, most of transposases belonged to IS50 families frequently located on plasmids, while the IS were mainly represented by IS30 and IS3 found in various bacteria and being responsible for information transfer and extreme adaptation. This fact suggests the high adaptability potential of \textit{L. pentosus} CF2-10N enabling the bacterium to withstand different environmental and gut stress conditions. Furthermore, the presence of eight prophage regions in the \textit{L. pentosus} CF2-10N genome highlights once more the genetic diversity and fitness of its genome, conferring a selective advantage for the survivability and resistance of this strain in view of the potential risk of losses associated with phage infection in different ecosystems. The presence of prophages in lactobacilli genomes is widely distributed (more than 92%, Sun et al., 2015) and is species-specific (Pei et al., 2021), while being highly dependent on the habitat. In this regard \textit{L. pentosus} CF2-10N contained intact lactobacilli prophage and incomplete or questionable prophage fragments similar to other bacteria (\textit{Staphylococcus}, \textit{Escherichia} and Enterobacteria phages) indicating its adaptability to harsh conditions (fermentation) which may confer flexibility against various stress triggers (phages from different sources such as air, water or soil). Other defense mechanisms were predicted in the \textit{L. pentosus} CF2-10N including a CRISPR system (CRISPR-I and CRISPR-II) represented by five CRISPR unquestionable arrays and 13 CRISPR associated proteins (six of them were exclusive of this strain) organized in two operons. This acquired immunity system, which provides protection against mobile genetic elements (conjugative plasmids, transposable elements, and phages) in \textit{L. pentosus} CF2-10N, was slightly different from \textit{L. pentosus} MP-10 isolated from the same ecological niche. Notably, 11 CRISPR associated proteins and 9 CRISPR arrays (3 of them were questionable CRISPRs) were detected in \textit{L. pentosus} MP-10, which indicated that the increased fitness greatly depends on the strain itself, under changing ecological lifestyles. Among the six newly detected genes, the CRISPR-I system was found to be coding for a Type II-A CRISPR-associated protein Cas2, involved in CRISPR adaptation for new spacer acquisition (Nam et al., 2011) and was associated with the \textit{cas9-cas1-cas2} cassette. Furthermore, the other genes (\textit{gene\_2925} [\textit{cas} 3] and a cascade of five genes coding for Type I-E CRISPR associated proteins) were found to be involved in interference and infection neutralization as reported by Xue and Sashital (2019).

Concerning functional properties, \textit{L. pentosus} CF2-10N genome analysis revealed the presence of genes coding for
adhesion, exopolysaccharide biosynthesis, tolerance to low pH and bile salts, immunomodulation, as well as vitamin and enzyme production. In this context, the adhesion capacity exhibited by this strain in vitro to Enterocyte-like Caco-2 ECACC86010202 (from colon adenocarcinoma) and HeLa 229 ECACC86090201 (from vaginal cervix carcinoma) cells (Pérez Montoro et al., 2016) was confirmed by the presence of genes coding for several adhesion/multifunctional proteins such as mucus-binding proteins, fibrinectin/fibrinogen-binding protein, Chitin-binding protein, ABC superfamily ATP binding cassette transporter, binding protein, cell surface proteins, manganese ABC transporter substrate-binding protein, elongation factor Tu, Molecular chaperone DnaK, molecular chaperone GroEL, co-chaperone GroES, class A sortase and type I glyceraldehyde-3-phosphate dehydrogenase. These proteins were reported to be involved in the adhesion to intestinal epithelial cells (Granato et al., 2004; Vélez et al., 2007; Lebeer et al., 2008; Sánchez et al., 2011; Jensen et al., 2014; Hymes et al., 2016), however, some of these proteins can also be involved in other functions such as stress response, drug efflux, carbohydrate transport and metabolism and other probiotic actions (Lebeer et al., 2008; Lewis et al., 2012; Monteagudo-Mera et al., 2019). The specific functionality notably depends on the surrounding conditions which induce gene expression, with differences detected in both in vitro and in vivo scenarios. On the other hand, other genes coding for proteins involved in cell recognition and adhesion to intestinal mucosae such as the four genes coding for exopolysaccharide biosynthesis proteins were identified in the L. pentosus CF2-10 N genome. These were found to be identical to those detected in L. pentosus MP-10 isolated from Aloreña table olives (Abriouel et al., 2016). Besides their role in niche adaptation, promoting auto-aggregation and biofilm formation, these proteins were also attributed anti-inflammatory, antioxidant, antiviral and antiproliferative activity functions through their interaction with the immune system (Castro-Bravo et al., 2018; Nguyen et al., 2020; Riaz Rajoka et al., 2020).

To allow the adaptation to different lifestyles, L. pentosus CF2-10 N harbored in its genome several genes involved in stress response such as acids and bile. In this sense, Pérez Montoro et al. (2016) reported the strain’s excellent tolerance properties in vitro (acidic pH of 1.5, up to 4% of bile salts and 5 mM of nitrate), while in the present study we detected for the first time several genes coding for proteins involved in bile/acidcs resistance particularly including cell protection (dnaK and groL), modifications in cell membranes (genes coding for Na+/H+ antiporter NhaC, lepA, pyrD), general function (genes coding for GNAT family acetyltransferase), and key components of central metabolism (pgk, gpm, CysK, luxS, tuf, efp, tif, FD24_GL002972, greA, greA2, fuaA) as it was reported elsewhere for other bacteria (Wu et al., 2010; Liu et al., 2018; Bagon et al., 2021). Most of these proteins are considered moonlighting proteins involved in adhesion to the intestinal epithelium among other functions (Pagnini et al., 2018).

Concerning attractive and promising biotechnological features revealed by in silico analysis of the L. pentosus CF2-10 N genome, detected enzymes were involved in the degradation of toxic/complex substrates such as tannase, alpha-amylase, amylopullulanase, beta-galactosidase, aminopeptidase, lipase esterase, peptidases, alpha/beta hydrolase, phenolic acid decarboxylase, carboxylesterase, alpha-acetolactate decarboxylase and multicopper oxidase. These findings indicate the high adaptability of this strain to a broad range of environmental niches, food matrices and also the gastrointestinal tract, while being able to ferment lactose and starch. Findings further demonstrate the strain’s potential ability to synthesize and degrade a broad array of simple and complex carbohydrates, such as starch, pullulan, amylopectin, tannin, beta-galactosides, phenolic acids and other substrates. It is further noteworthy that L. pentosus CF2-10 N harbored genes coding for vitamin biosynthesis such as the vitamin B group (B1 or thiamine, B2 or riboflavin and B5), folate and vitamin K2 or menaquinone. In this regard, preliminary in vitro studies hinted towards a potential vitamin production ability of L. pentosus CF2-10 N. However, future studies are necessary and will be performed to investigate this potential in further detail.

Conclusion

The results obtained in the present study support the hypothesis that L. pentosus CF2-10 N is an excellent probiotic candidate of vegetable origin. Notably, besides fulfilling the main criteria for probiotic selection in vitro as shown by our previous studies, in silico genome analysis in this study revealed novel insights into its safety and functionality, greatly highlighting the microorganism’s ecological flexibility and adaptability to a broad range of environmental niches, food matrices and the gastrointestinal tract. The safety of L. pentosus CF2-10 N was further confirmed by the absence of virulence determinants and acquired antibiotic resistance genes, with the resistome mostly represented by efflux-pump resistance genes responsible for the intrinsic resistance exhibited by this strain. On the other hand, defense mechanisms of L. pentosus CF2-10 N consist of eight prophage regions as well as a CRISPR (clustered regularly interspaced short palindromic repeats)/cas (CRISPR-associated protein genes) system (CRISPR-I and CRISPR-II) as acquired immune system against mobile elements. The latter is notably represented by five CRISPR unquestionable arrays and 13 CRISPR associated proteins (six of them were exclusive of this strain). Furthermore, the functionality of this strain was supported by the presence of genes coding for proteins involved in adhesion, exopolysaccharide biosynthesis, tolerance to low pH and bile salts, immunomodulation as well as vitamin and enzyme production.

Taken together these results we suggest that L. pentosus CF2-10 N could be considered as potential and promising probiotic candidate able to colonize several niches and adapt to different lifestyles, while providing attractive probiotic features, which will be explored in vivo in future studies with the aim to be applied in vegetable fermentations (including olives) and/or other substrates.
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

HA and NB conceived, designed the experiments, and drafted the paper. HA, JM, NC, and NB performed the experiments and analyzed the data. HA contributed reagents, materials, and analysis tools. All authors contributed to the article and approved the submitted version.

Acknowledgments

We acknowledge the Research Team (University of Jaen, El_BIO1_2021).

References

Abriouel, H., Benomar, N., Cobos, A., Caballero, N., Fernández Fuentes, M. A., Pérez-Pulido, R., et al. (2012). Characterization of lactic acid bacteria from naturally-fermented Manzanilla Aloreña green table olives. Food Microbiol. 32, 308–316. doi: 10.1016/j.fm.2012.07.006

Abriouel, H., Benomar, N., Lucas, R., and Gálvez, A. (2011). Culture-independent study of the diversity of microbial populations in brines during fermentation of naturally-fermented Aloreña green table olives. Int. J. Food Microbiol. 144, 487–496. doi: 10.1016/j.ijfoodmicro.2010.11.006

Abriouel, H., Pérez Montoro, B., Casado Muñoz, M. D. C., Lavilla Lerma, L., Hidalgo Pestaña, M., Caballero Gómez, N., et al. (2016). Complete genome sequence of a potential probiotic, lactobacillus pentosus MP-10, isolated from fermented Aloreña table olives. Genome Announc. 4, e00854–e00816. doi: 10.1128/genomeA.00854-16

Abriouel, H., Pérez Montoro, B., Casimiro-Soriguero, C. S., Pérez Pulido, A. J., Knapp, C. W., Caballero Gómez, N., et al. (2017). Insight into potential probiotic markers predicted in lactobacillus pentosus MP-10 genome sequence. Front. Microbiol. 8:891. doi: 10.3389/fmicb.2017.00891

Abriouel, H., Pérez Montoro, B., de la Fuente Ordoñez, J. J., Lavilla Lerma, L., Knapp, C. W., and Benomar, N. (2019). New insights into the role of plasmids from probiotic lactobacillus pentosus MP-10 in Aloreña table olive brine fermentation. Sci. Reports 9:10938. doi: 10.1038/s41598-019-47384-1

Aloksh, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., et al. (2020). CARD 2020: antibiotic resistance surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res. 48, D517–D525. doi: 10.1093/nar/gkz935

Allain, T., Chaouch, S., Thomas, M., Vallée, I., Buret, A. G., Langella, P., et al. (2018). Bile-salt-hydrolyses from the probiotic strain lactobacillus johnsonii L1, mediated anti-giardial activity in vitro and in vivo. Front. Microbiol. 8:2707. doi: 10.3389/fmicb.2017.02707

Alonso García, A., Benomar, B., Lavilla Lerma, F., de la Fuente Ordoñez, J. J., Knapp, C. W., and Abriouel, H. (2023). Changes in resistance profile of potential probiotic Lactiplantibacillus pentosus in response to edible oil adaptation. Food Microbiol. 104:1418. doi: 10.1016/j.fm.2022.104148

Alonso García, E., de la Fuente Ordoñez, J. J., Lavilla Lerma, L., Estudillo-Martínez, M. D., Castillo-Gutiérrez, S., Benomar, N., et al. (2021). Transcriptomic profile and probiotic properties of Lactiplantibacillus pentosus pre-adapted to edible oils. Front. Microbiol. 12:747043. doi: 10.3389/fmicb.2021.747043

Amara, A. A., and Shibil, A. (2015). Role of probiotics in health promotion, infection control and disease treatment and management. Saudi Pharm. J. 23, 107–114. doi: 10.1016/j.jspj.2013.07.001

Anukam, K. C., Macklaim, J. M., Gloor, G. B., Reid, G., Boekhorst, J., Renckens, B., et al. (2013). Genome sequence of lactobacillus pentosus KCA1, vaginal isolate from a healthy premenopausal woman. PLoS One 8:e59239. doi: 10.1371/journal.pone.0059239

Arndt, D., Grant, J., Marcus, 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

Bagos, B. B., Valeriano, V. D. V., Oh, J. K., Pajarillo, E. A. B., Lee, J. Y., and Kang, D. K. (2021). Exoproteome perspective on the bile stress response of lactobacillus johnsonii. Proteomes 9:10. doi: 10.3390/proteomes9010010

Benítez-Cabello, A., Calero-Delgado, B., Rodríguez-Gómez, F., Garrido-Fernández, A., Jiménez-Díaz, R., and Arroyo-López, F. N. (2019). Biodiversity and multifunctional features of lactic acid bacteria isolated from table olive biofilms. Food Microbiol. 10.1016/j.fm.2018.06.020

Biswas, A., Staals, R. H. J., Morales, S. E., Fineran, P. C., and Brown, C. M. (2016). CRISPRDetect: a flexible algorithm to define CRISPR arrays. BMC Genom. 17:356. doi: 10.1186/s12864-016-2627-0

Bortolaiia, V., Kaas, R. S., Berriman, M., Rajandream, M. A., Barrell, B. G., and Parkhill, J. (2005). ACT: the Artemis comparison tool. Nucleic Acids Res. 33, 3877–3886. doi: 10.1093/nar/gki258

Brady, C. W.,, Caballero Gómez, N., et al. (2017). Insight into potential probiotic markers predicted in lactobacillus pentosus MP-10 genome sequence. Front. Microbiol. 8:891. doi: 10.3389/fmicb.2017.00891

Carrasco, J. A., Lucena-Ordóñez, H., Brenes, M., and Ruiz-Barba, J. I. (2018). Expression of genes involved in metabolism of phenolic compounds by lactobacillus pentosus and its relevance for table-olive fermentations. Food Microbiol. 76, 382–389. doi: 10.1016/j.fm.2018.06.020

Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M. A., Barrell, B. G., and Parkhill, J. (2005). ACT: the Artemis comparison tool. Bioinform. 21, 3422–3423. doi: 10.1093/bioinformatics/bti553
Carver, T. J., Thomson, N., Bleasby, A., Berriman, M., and Parkhill, J. (2009). DNAPlotter: circular and linear interactive genome visualization. Bioinfrom. 25, 119–120. doi: 10.1093/bioinformatics/btn578

Casado Muñoz, M. C., Benomar, N., Ennahar, S., Horvatovitch, P., Lavilla Lerma, L., Knapp, C. W., et al. (2016). Comparative proteomic analysis of a potentially probiotic lactobacillus plantarum MP-10 for the identification of key proteins involved in antibiotic resistance and biodegradation tolerance. Int. J. Food Microbiol. 222, 8–15. doi: 10.1016/j.ijfoodmicro.2016.01.012

Casado Muñoz, M. C., Benomar, N., Lerma, I. L., Galvez, A., and Abriouel, H. (2014). Functional resistance of lactobacillus plantarum Leuconostoc pseudomesenteroides isolated from naturally-fertilized Aloreña table olives throughout fermentation process. Int. J. Food Microbiol. 172, 110–118. doi: 10.1016/j.ijfoodmicro.2013.11.025

Castro-Bravo, N., Wells, J. M., Margolles, A., and Ruas-Madiedo, P. (2018). Interactions of surface exopolysaccharides from Bifidobacterium and lactobacillus within the intestinal environment. Front. Microbiol. 9:2426. doi: 10.3389/fmicb.2018.02426

Chin, C. S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Hein, L. E., et al. (2005). Small read-set diversity of conserved genomic sequence with rearrangements. Bioinfrom. 21, 3674–3676. doi: 10.1093/bioinformatics/bti1610

Darling, A. C. E., Mau, B., Blattner, F. R., and Perna, N. T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 14, 1396–1406. doi: 10.1101/gr.2289704

Duaz, R. M., Lin, X. B., Zheng, J. Z., Martino, M. E., Grenier, T., Pérez-Muñoz, M. E., et al. (2017). Lifeslstyles in transition: evolution and natural history of the genus lactobacillus. FEMS Microbiol. Res. 41, S27–S58. doi: 10.1093/femsec/ffx039

Granato, D., Bergonzelli, G. E., Priddmore, R. D., Marvin, L., Rouvet, M., and Cortiere, Thibaut (2004). Cell-surface-associated transport factor Tscus and the attachment of lactobacillus johnsonii NCC533 (Lai) to human intestinal cells and mucins. Infect. Immun. 72, 2160–2169. doi: 10.1128/IAI.72.4.2160-2169.2004

Guantario, B., Zimno, P., Schifano, E., Rosselli, M., Perozzi, G., Palleschi, C., et al. (2018). In vitro and in vivo selection of potentially probiotic lactobacilli from Nocellara del Belice table olives. Front. Microbiol. 9:595. doi: 10.3389/fmicb.2018.00595

Hill, C., Guarner, F., Reid, G., Gibson, G. R., Merenstein, D. J., Pot, B., et al. (2014). The international scientific association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. Nat. Rev. Gastroenterol. Hepatol. 11, 506–514. doi: 10.1038/nrgastro.2014.66

Hymes, J. P., Johnson, B. R., Barrangou, R., and Klaenhammer, T. R. (2016). Functional analysis of an S-layer-associated fibronectin-binding protein in Lactobacillus acidophilus NCFM. Appl. Environ. Microbiol. 82, 2676–2685. doi: 10.1128/AEM.00214-16

Jensen, H., Roos, S., Jonsson, H., Rud, I., Grimmer, C., and Hynes, J. P. (2019). tRNAscan-SE: a program for improved annotation, visualization and analysis in functional genomics research. Bioinformatics 35, 4233–4240. doi: 10.1093/bioinformatics/btz318

Johannsmeier, S. D., Arnekell, K., Pagan-Medina, C. G., Méndez-Sandoval, L., Arellano, C., et al. (2021). Genotypic and phenotypic diversity among lactobacillus plantarum and lactobacillus pentosus isolated from industrial scale cucumber fermentations. Front. Microbiol. 12:63562. doi: 10.3389/fmicb.2021.63562

Perpetuini, G., Perete, R., García-González, N., Chávez, C. A., and Cortesi, A. (2016). Mucosal adhesion and anti-inflammatory effects of lactobacillus reuteri isolated from Spanish-style green olive fermentations. FEMS Microbiol. Lett. 369, 1123–1126. doi: 10.1093/femsle/fnq056

Nguyen, P. T., Nguyen, T. T., Bui, D. C., Hong, P. T., Hoang, Q. K., and Nguyen, H. T. (2020). Exopolysaccharide production by lactic acid bacteria: the manipulation of environmental stresses for industrial applications. AIMS Microbiol. 6, 451–469. doi: 10.3934/microbiol.2020027

Oguntubayo, F. A., and Nair, R. B. (2015). Multifunctional properties of lactobacillus plantarum strains isolated from fermented cereal foods. Funct. Foods 7, 621–631. doi: 10.1016/j.foodbiopro.2015.06.022

Pagnini, C., Coreto, V. D., Martorelli, M., Laini, C., D’Ambra, G., Di Giulio, E., et al. (2018). Mucosal adhesion and anti-inflammatory effects of lactobacillus rhamnosus GG in the human colonic mucosa: a proof-of-concept study. World J. Gastroenterol. 24, 4652–4662. doi: 10.3748/wjg.v24.i41.4652

Pérez-Montoro, B., Benomar, N., Caballero Gómez, N., Ennahar, S., Horvatovitch, P., Knapp, C. W., et al. (2018a). Proteomic analysis of lactobacillus pentosus for the identification of potential markers of adhesion and other probiotic features. Food Res. Int. 111, 58–66. doi: 10.1016/j.foodres.2018.04.072

Pérez-Montoro, B., Benomar, N., Caballero Gómez, N., Ennahar, S., Horvatovitch, P., Knapp, C. W., et al. (2018b). Proteomic analysis of lactobacillus johnsonii NCC533 (Lai) to human intestinal cells and mucins. Infect. Immun. 72, 2160–2169. doi: 10.1128/IAI.72.4.2160-2169.2004

Pérez-Montoro, B., Benomar, N., Caballero Gómez, N., Ennahar, S., Horvatovitch, P., Knapp, C. W., et al. (2018a). Proteomic analysis of lactobacillus johnsonii NCC533 (Lai) to human intestinal cells and mucins. Infect. Immun. 72, 2160–2169. doi: 10.1128/IAI.72.4.2160-2169.2004

Perretuini, G., Prete, R., García-González, N., Chávez, C. A., and Cortesi, A. (2016). Mucosal adhesion and anti-inflammatory effects of lactobacillus reuteri isolated from Spanish-style green olive fermentations. FEMS Microbiol. Lett. 369, 1123–1126. doi: 10.1093/femsle/fnq056

Riaz Rajoka, M. S., Wu, Y., Mehwish, H. M., Bansal, M., and Zhao, L. (2020). Lactobacillus exopolysaccharides: new perspectives on engineering strategies, physiochemical functions, and immunomodulatory effects on host health, trends foods Sci. Technol. 103, 36–48. doi: 10.1016/j.tifs.2020.06.003

Rodríguez-Gómez, F., Romero-Gil, V., Arroyo-López, F. N., Roldán-Reyes, J. C., Torres-Gallardo, R., Bautista-Gallego, J., et al. (2017). Assessing the challenges in the manipulation of environmental stresses for industrial applications. Food Microbiol. 72, 31–38. doi: 10.1016/j.foodmicro.2017.11.006

Sánchez, B., González-Tejedo, C., Ruas-Madiedo, P., Urdaci, M. C., and Margolles, A. (2011). Antibiotic resistance of lactobacillus plantarum and lactobacillus pentosus strains isolated from Spanish-style green olive fermentations. Food Microbiol. 28, 2196–2204. doi: 10.1016/j.tifs.2020.06.003

Sánchez, B., González-Tejedo, C., Ruas-Madiedo, P., Urdaci, M. C., and Margolles, A. (2011). Antibiotic resistance of lactobacillus plantarum and lactobacillus pentosus strains isolated from Spanish-style green olive fermentations. Food Microbiol. 28, 2196–2204. doi: 10.1016/j.tifs.2020.06.003

Abriouel et al. 10.3389/fmicb.2022.989824
with anti-proliferative activity induce cytokine/chemokine production and neutrophil recruitment in mice. *Benef. Microbes* 8, 615–623. doi: 10.3389/fmicb.2022.989824

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinform.* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Sun, Z., Harris, H. M., McCann, A., Guo, C., Argimon, S., Zhang, W., et al. (2015). Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera. *Nat. Commun.* 6:8322. doi: 10.1038/ncomms9322

Tofalo, R., Perpetuini, G., Schirone, M., Ciarrocchi, A., Fasoli, G., Suzzi, G., et al. (2014). *Lactobacillus pentosus* dominates spontaneous fermentation of Italian table olives. *LWT-Food Sci. Technol.* 57, 710–717. doi: 10.1016/j.lwt.2014.01.035

Vaccalluzzo, A., Pino, A., Russo, N., De Angelis, M., Caggia, C., and Randazzo, C. L. (2020). FoodOmics as a new frontier to reveal microbial community and metabolic processes occurring on table olives fermentation. *Food Microbiol.* 82:103606. doi: 10.1016/j.fm.2020.103606

van de Wijgert, J. H. H. M., and Verwijs, M. C. (2019). Lactobacilli-containing vaginal probiotics to cure or prevent bacterial or fungal vaginal dysbiosis: a systematic review and recommendations for future trial designs. *BJOG* 127, 287–299. doi: 10.1111/1471-0528.15870

Vélez, M. P., De Keersmaecker, S. C., and Vanderleyden, J. (2007). Adherence factors of *lactobacillus* in the human gastrointestinal tract. *FEMS Microbiol. Lett.* 276, 140–148. doi: 10.1111/j.1574-6968.2007.00908.x

Wu, R., Sun, Z., Wu, J., Meng, H., and Zhang, H. (2010). Effect of bile salts stress on protein synthesis of *lactobacillus casei* Zhang revealed by 2-dimensional gel electrophoresis. *J. Dairy Sci.* 93, 3858–3868. doi: 10.3168/jds.2009-2967

Xue, C., and Sashital, D. G. (2019). Mechanisms of type I-E and I-F CRISPR-Cas Systems in *Enterobacteriaceae*. *EcoSal Plus* 8:10.1128/ecosalplus.ESP-0008-2018. doi: 10.1128/ecosalplus.ESP-0008-2018

Yan, S., Tian, Z., Li, M., Li, B., and Cui, W. (2019). Effects of probiotic supplementation on the regulation of blood lipid levels in overweight or obese subjects: a meta-analysis. *Food Funct.* 10, 1747–1759. doi: 10.1039/c8fo02163e

Yoha, K. S., Nida, S., Dutta, S., Moses, J. A., and Anandharamakrishnan, C. (2022). Targeted delivery of probiotics: perspectives on research and commercialization. *Probiotics Antimicro. Prot.* 14, 15–48. doi: 10.1007/s12602-021-09791-7

Zankari, E., Hasman, H., Costerino, S., Vestergaard, M., Rasmussen, S., Lund, O., et al. (2012). Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* 67, 2640–2644. doi: 10.1093/jac/dks261

Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7, 203–214.

Zheng, J., Wittouck, S., Salvetti, E., Franz, C. M. A. P., Harris, H. M. B., Mattarelli, P., et al. (2020). A taxonomic note on the genus *lactobacillus* description of 23 novel genera, emended description of the genus *lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. *Int. J. Syst. Evol. Microbiol.* 70, 2782–2858. doi: 10.1099/ijsem.0.004107

Zhou, Y., Liang, Y., Lynch, K., Dennis, J. J., and Wishart, D. S. (2011). PHAST: a fast phage search tool. *Nucl. Acids Res.* 39, W347–W352. doi: 10.1093/nar/gkr485

Zotta, T., Giavalisco, M., Parente, E., Picariello, G., Stano, F., and Ricciardi, A. (2022). Selection of *Lactiplantibacillus* strains for the production of fermented table olives. *Microorganisms* 10:625. doi: 10.3390/microorganisms10030625