



Comparative Genomics of *Lactobacillus acidipiscis* ACA-DC 1533 Isolated From Traditional Greek Kopanisti Cheese Against Species Within the *Lactobacillus salivarius* Clade

Maria Kazou¹, Voula Alexandraki¹, Jochen Blom², Bruno Pot³, Effie Tsakalidou¹ and Konstantinos Papadimitriou^{1*}

¹ Laboratory of Dairy Research, Department of Food Science and Human Nutrition, Agricultural University of Athens, Athens, Greece, ² Bioinformatics and Systems Biology, Justus-Liebig-University Giessen, Giessen, Germany, ³ Research Group of Industrial Microbiology and Food Biotechnology (IMDO), Department of Bioengineering Sciences (DBIT), Vrije Universiteit Brussel, Brussels, Belgium

OPEN ACCESS

Edited by:

Frank T. Robb,
University of Maryland, Baltimore,
United States

Reviewed by:

Jinshui Zheng,
Huazhong Agricultural University,
China
Seong Woon Roh,
World Institute of Kimchi (WIKIM),
South Korea
Sergei Kozyavkin,
Fidelity Systems, United States

*Correspondence:

Konstantinos Papadimitriou
kpapadimitriou@aua.gr

Specialty section:

This article was submitted to
Evolutionary and Genomic
Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 25 October 2017

Accepted: 23 May 2018

Published: 11 June 2018

Citation:

Kazou M, Alexandraki V, Blom J,
Pot B, Tsakalidou E and
Papadimitriou K (2018) Comparative
Genomics of *Lactobacillus acidipiscis*
ACA-DC 1533 Isolated From
Traditional Greek Kopanisti Cheese
Against Species Within
the *Lactobacillus salivarius* Clade.
Front. Microbiol. 9:1244.
doi: 10.3389/fmicb.2018.01244

Lactobacillus acidipiscis belongs to the *Lactobacillus salivarius* clade and it is found in a variety of fermented foods. Strain ACA-DC 1533 was isolated from traditional Greek Kopanisti cheese and among the available *L. acidipiscis* genomes it is the only one with a fully sequenced chromosome. *L. acidipiscis* strains exhibited a high degree of conservation at the genome level. Investigation of the distribution of prophages and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs) among the three strains suggests the potential existence of lineages within the species. Based on the presence/absence patterns of these genomic traits, strain ACA-DC 1533 seems to be more related to strain JCM 10692^T than strain KCTC 13900. Interestingly, strains ACA-DC 1533 and JCM 10692^T which lack CRISPRs, carry two similar prophages. In contrast, strain KCTC 13900 seems to have acquired immunity to these prophages according to the sequences of spacers in its CRISPRs. Nonetheless, strain KCTC 13900 has a prophage that is absent from strains ACA-DC 1533 and JCM 10692^T. Furthermore, comparative genomic analysis was performed among *L. acidipiscis* ACA-DC 1533, *L. salivarius* UCC118 and *Lactobacillus ruminis* ATCC 27782. The chromosomes of the three species lack long-range synteny. Important differences were also determined in the number of glycozyme related proteins, proteolytic enzymes, transporters, insertion sequences and regulatory proteins. Moreover, no obvious genomic traits supporting a probiotic potential of *L. acidipiscis* ACA-DC 1533 were detected when compared to the probiotic *L. salivarius* UCC118. However, the existence of more than one glycine-betaine transporter within the genome of ACA-DC 1533 may explain the ability of *L. acidipiscis* to grow in fermented foods containing high salt concentrations. Finally, *in silico* analysis of the *L. acidipiscis* ACA-DC 1533 genome revealed pathways that could underpin the production of major volatile compounds during the catabolism of amino acids that may contribute to the typical piquant flavors of Kopanisti cheese.

Keywords: *Lactobacillus*, genome, pseudogene, motility, horizontal gene transfer, phage, probiotic, metabolism

INTRODUCTION

The genus *Lactobacillus* constitutes a diverse group of bacteria comprising more than 200 species and subspecies¹ that are ubiquitous and frequently found in a variety of nutrient-rich ecological niches (Pot et al., 2014; Sun Z. et al., 2015). Lactobacilli produce lactic acid as the main end-product of carbohydrate fermentation allowing them to prevail in microbial ecosystems. This attribute along with their safety profile and their ability to shape organoleptic characteristics of the final product are the central reasons for their extensive use in artisanal or industrial food fermentations (Bernardeau et al., 2008; Sun Z. et al., 2015; Reginensi et al., 2016). Apart from food-related lactobacilli, the genus includes many commensals of the human, animal and plant microbiota (Cannon et al., 2005; Duar et al., 2017). The available genomes for *Lactobacillus* species and the close phylogenetic relationship among food- and host-related strains offer a wealth of information that underpin specialized mechanisms of bacterial adaptation to different environments (Sun Z. et al., 2015).

Lactobacillus acidipiscis is a salt-tolerant species originally isolated from fermented fish (Tanasupawat et al., 2000). The species has been also found in a variety of cheeses, i.e., Halloumi (Lawson et al., 2001; Naser et al., 2006; Kim et al., 2011), Cotija (Morales et al., 2011), Minas (Perin et al., 2017), and double cream cheese (Morales et al., 2011; Melgar-Lalanne et al., 2013), as well as in fermented fish (An et al., 2010; Tsuda et al., 2012; Thamacharoensuk et al., 2017), fermented meat (Nguyen et al., 2013), sake (Koyanagi et al., 2016), pickles (Arasu et al., 2015), grasses (Tohno et al., 2012; Khota et al., 2016), mulberry silage (Altaher et al., 2015), sweet paste (Mao et al., 2017), the traditional Chinese fermented vegetable Sichuan paocai, (Cao et al., 2017) and table olives (Randazzo et al., 2017). Moreover, *L. acidipiscis* has also been found in vinegar (Li P. et al., 2014) and soy sauce, where it is considered to be a spoiler (Tanasupawat et al., 2002; Cheng et al., 2014; Li N. et al., 2014). *L. acidipiscis* ACA-DC 1533 was isolated from traditional Greek Kopanisti cheese and along with *Lactobacillus rennini* were the dominant microbiota of this cheese (Asteri et al., 2009). Interestingly, both of them produced alcohols and carbonyl compounds most probably *via* amino acid catabolism that may contribute to the typical piquant

flavors of Kopanisti cheese (Yvon and Rijnen, 2001; Asteri et al., 2009; Donnelly, 2016).

Phylogenetic analysis of *L. acidipiscis* places the bacterium in the *Lactobacillus salivarius* clade. The *L. salivarius* clade is the second largest group of lactobacilli with 27 recognized species following that of *Lactobacillus delbrueckii* (29 species; Pot et al., 2014). The *L. salivarius* clade consists mainly of commensal isolates and to a lesser degree of strains found in fermented foods (Cousin et al., 2015). Several strains belonging to the clade exhibit putative probiotic traits (Neville and O'Toole, 2010). Therefore, comparative genomics among members of the *L. salivarius* clade may reveal important aspects, such as niche adaptation, technological potential, and probiotic properties (Forde et al., 2011; Raftis et al., 2011; Sun Z. et al., 2015). So far, there are eight genomes with fully sequenced chromosomes in the *L. salivarius* clade publicly available in the NCBI database, i.e., six from *L. salivarius* (Claesson et al., 2006; Jimenez et al., 2010; Raftis et al., 2014; Chenoll et al., 2016), one from *Lactobacillus ruminis* (Forde et al., 2011) and one from *L. acidipiscis* (Kazou et al., 2017). Furthermore, *L. acidipiscis* JCM 10692^T and DSM 15836^T isolated from fermented fish as well as *L. acidipiscis* DSM 15353 and KCTC 13900 isolated from Halloumi cheese have been partially sequenced (Kim et al., 2011; Sun Z. et al., 2015). In fact, strains JCM 10692^T and DSM 15836^T are replicas of the same strain^{2,3} and the same applies for strains DSM 15353 and KCTC 13900^{4,5}.

The genome sequence of *L. acidipiscis* ACA-DC 1533 has been published (Kazou et al., 2017) and the current study aims to examine aspects of the evolution, physiology, metabolism and technological properties of the species according to the available *L. acidipiscis* genomes. Furthermore, we perform comparative genomics among the species with fully sequenced genomes in the *L. salivarius* clade to shed light to niche adaptation (host or food related, or both). Our analysis reveals technological properties of *L. acidipiscis* ACA-DC 1533 that may support the potential use of the isolate in food fermentations.

MATERIALS AND METHODS

Chromosome-Plasmid Sequences and Annotations

Species/strains employed in phylogenetic analysis and comparative genomics are presented in **Supplementary Table S1**. All annotated sequences derived from RefSeq version 86 with the exception of plasmids pLAC2 and pLAC3 of *L. acidipiscis* ACA-DC 1533 that have not been included in RefSeq yet, so we used their GenBank/ENA versions (Kazou et al., 2017). In the table we present all relevant information to aid the reader assess whether differences or similarities in gene content among strains analyzed may be influenced by differences in sequencing

¹<http://www.bacterio.net/lactobacillus.html>

Abbreviations: ABC, ATP-binding cassette; ANI, average nucleotide identity; CAZy, Carbohydrate Active Enzymes; CBMs, Carbohydrate-Binding Modules; CEs, Carbohydrate Esterases; COG, Clusters of Orthologous Groups; CRISPRs, Clustered Regularly Interspaced Short Palindromic Repeats; dBBQs, dataBase of Bacterial Quality scores; dbCAN, DataBase for automated Carbohydrate-active enzyme Annotation; EPS, Exopolysaccharide; GHs, Glycoside Hydrolases; GIs, Genomic islands; GTs, Glycosyl Transferases; HGT, Horizontal Gene Transfer; HK, Histidine Kinase; ISs, insertion sequences; KEGG, Kyoto Encyclopedia of Genes and Genomes; LCBS, locally collinear blocks; MFS, major facilitator superfamily; OCSs, One-component Systems; ODPs, Other DNA-binding Proteins; PBMCs, peripheral blood mononuclear cells; PEP-PTS, Phosphoenolpyruvate Phosphotransferase System; P2RP, Predicted Prokaryotic Regulatory Proteins; r2cat, Related Reference Contig Arrangement; RM, Restriction-Modification; RPs, Regulatory Proteins; RR, Response Regulator; RSM, reconstituted skim milk; SFs, Sigma Factors; TA, toxin-antitoxin; TCSs, Two-component Systems; TFs, Transcription Factors; TRs, Transcriptional Regulators.

²http://www.jcm.riken.jp/cgi-bin/jcm/jcm_number?JCM=10692

³<https://www.dsmz.de/catalogues/details/culture/DSM-15836.html>

⁴<https://www.dsmz.de/catalogues/details/culture/DSM-15353.html>

⁵http://kctc.kribb.re.kr/En/jsearch/j_sview.aspx?sn=13900

technologies and/or tools used for sequence assembly and annotation.

Phylogenetic Analysis

A whole genome phylogenetic tree based on the core genes among representative strains of all species in the *L. salivarius* clade using *L. acidipiscis* ACA-DC 1533 as the reference genome was constructed with the EDGAR software (Blom et al., 2009). It should be noted that whenever available, sequences of type strains were preferred. Core gene sets were aligned using MUSCLE, the individual alignments were concatenated and the resulting genome alignment was used as input for the construction of the phylogenetic tree with the neighbor-joining method as implemented in the PHYLIP package. *Weissella kandleri* DSM 20593^T and *Lactobacillus delbrueckii* subsp. *bulgaricus* ATCC 11842^T were used as outgroups.

Comparative Genomic Analysis

To confirm the clonal relation among sequenced strains of *L. acidipiscis* as these are deduced from different databases, we used an ANI heat map as calculated with the EDGAR tool. The completeness of partial genome sequences of *L. acidipiscis* strains was assessed using the dBBQs (Wanchai et al., 2017). Preliminary evaluation of the presence of plasmids in the partially sequenced *L. acidipiscis* strains was performed with the r2cat tool (Husemann and Stoye, 2010), using as templates the three pLAC plasmid sequences of strain ACA-DC 1533. The circular map of *L. acidipiscis* ACA-DC 1533 was constructed by the DNAPlotter software (Carver et al., 2009). Pan/core-genome and singleton analysis were conducted with EDGAR. Comparison of the motility gene clusters among *L. acidipiscis* ACA-DC 1533 and KCTC 13900 as well as *Lactobacillus curvatus* NRIC 0822 was performed with the Easyfig comparison tool (Sullivan et al., 2011). The GenBank accession numbers for the motility operons of *L. acidipiscis* KCTC 13900 and *L. curvatus* NRIC 0822 are KM886858 and KM886863, respectively (Cousin et al., 2015). The EggNOG server version 4.5 was used for COG annotation (Huerta-Cepas et al., 2016). COG frequency heat maps with double hierarchical clustering were generated using the RStudio and the package “gplots”⁶. GIs, ISs, putative prophages, CRISPRs, RM systems, TA systems and putative antimicrobial peptides were predicted using the IslandViewer 4 web-based resource (Bertelli et al., 2017), the ISSaga platform (Varani et al., 2011), the PHASTER web server (Arndt et al., 2016), the CRISPRFinder web tool (Grissa et al., 2007), the REBASE database (Roberts et al., 2015), the TAFinder (Xie et al., 2018) and the BAGEL (van Heel et al., 2013), respectively. The glyco biome profile was investigated using the dbCAN (Yin et al., 2012) against the CAZy database (Lombard et al., 2014). Furthermore, transporters were determined using the TransportDB database (Elbourne et al., 2017). Pathways were assigned with the KEGG database (Kanehisa et al., 2016). Regulatory proteins including TCSs, TFs, and ODPs were detected with the P2RP web server (Barakat et al., 2013). Full-length chromosome alignments were

created by progressiveMAUVE (Darling et al., 2010). Finally, the carbohydrate fermentation profile of *L. acidipiscis* ACA-DC 1533 was determined using API 50 CHL stripes (bioMérieux, Marcy-l’Étoile, France).

RESULTS AND DISCUSSION

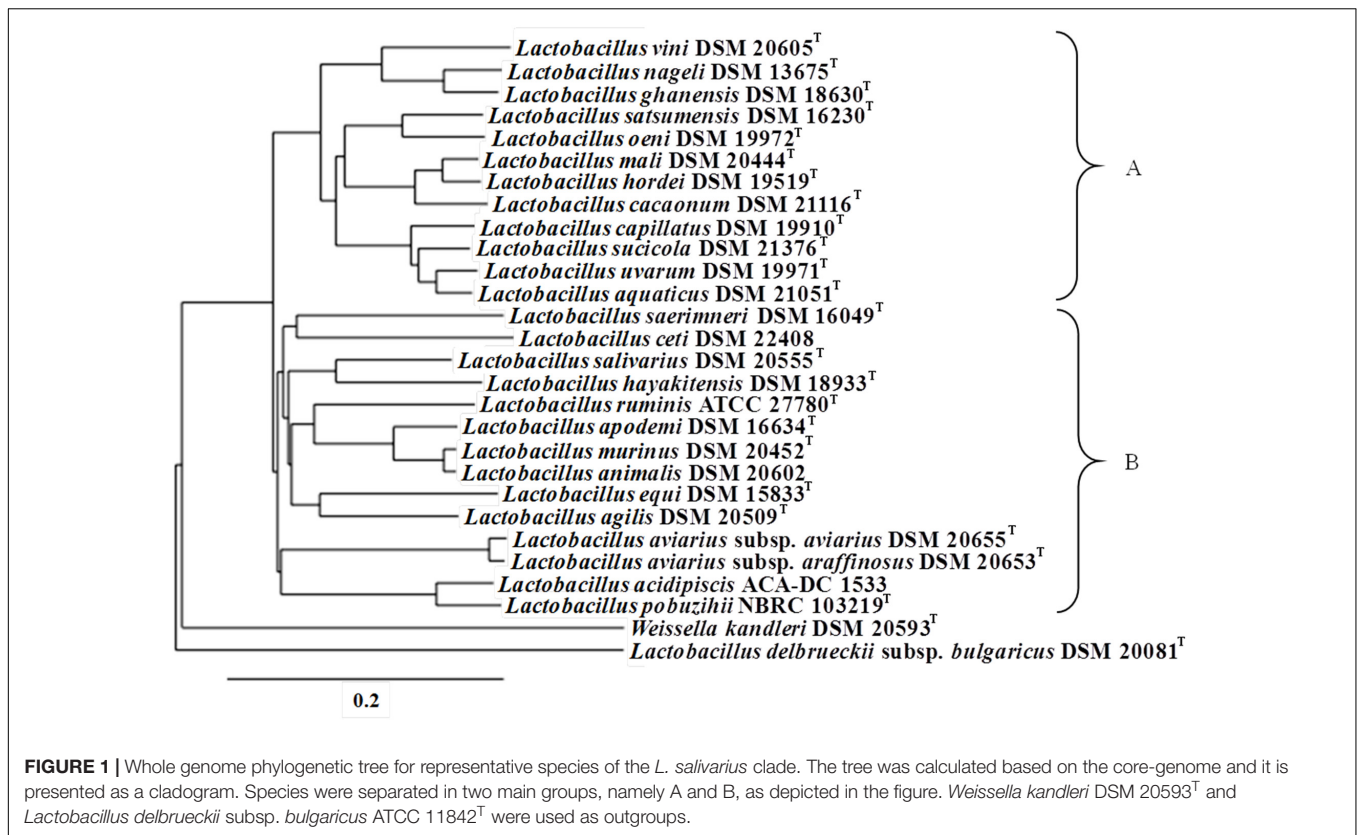
Whole Genome Phylogeny of the *L. salivarius* Clade

The phylogenetic relationship among the species of the *L. salivarius* clade was determined based on whole genome sequences. Analysis with the EDGAR software revealed two major clusters containing 12 and 14 species, respectively (Figure 1). *L. acidipiscis* was grouped together with *Lactobacillus pobuzihii* in a cluster, which also included *L. salivarius*. The strains employed in the phylogenetic analysis of the *L. salivarius* clade exhibited a pan genome of 13,470 genes, while the core-genome consisted of 349 genes. Moreover, proteins of the species belonging to the *L. salivarius* clade were distributed into various COG functional categories with a relatively distinct profile for each species. Interestingly, hierarchical clustering of the COG frequency heat map (Figure 2) revealed two clusters, which were very similar to the two clusters mentioned above that were obtained in the whole genome phylogenetic tree (Figure 1). It should be noted that *L. acidipiscis* ACA-DC 1533 was placed separately from these two clusters most probably due to an increased percentage of genes in the replication, recombination and repair (L) COG category. This difference could arise from a higher number of transposases in the ACA-DC 1533 genome but the number of transposases in the partial genomes employed during this analysis may be severely skewed. Nevertheless, *L. acidipiscis* also exhibited a higher number of transposases when compared to the complete genome sequences of *L. salivarius* and *L. ruminis* (please see below). Both whole genome phylogeny and COG analysis can be influenced by the partial nature of some of the sequences employed as well as differences in pipelines used for genome assembly and annotation. However, the whole genome phylogenetic tree is similar in the overall topology to the 16S rRNA phylogenetic tree of the entire *Lactobacillus* genus published by Pot et al. (2014) which is independent of genome completeness and annotation. The same applies when we compared our whole genome phylogenetic tree to the tree based on the concatenated amino acid sequences of 16 marker genes published by Sun Z. et al. (2015).

General Genomic Features of *L. acidipiscis* Strains

To date, there are five sequenced strains of *L. acidipiscis*, i.e., ACA-DC 1533, KCTC 13900, DSM 15353, JCM 10692^T and DSM 15836^T. As mentioned above, strains KCTC 13900 and DSM 15353 as well as JCM 10692^T and DSM 15836^T are replicas. Since this is not always obvious in the respective literature (Kim et al., 2011; Sun Z. et al., 2015), the relatedness among the two pairs of *L. acidipiscis* strains was also obtained by the ANI performed with

⁶<http://www.rstudio.org>

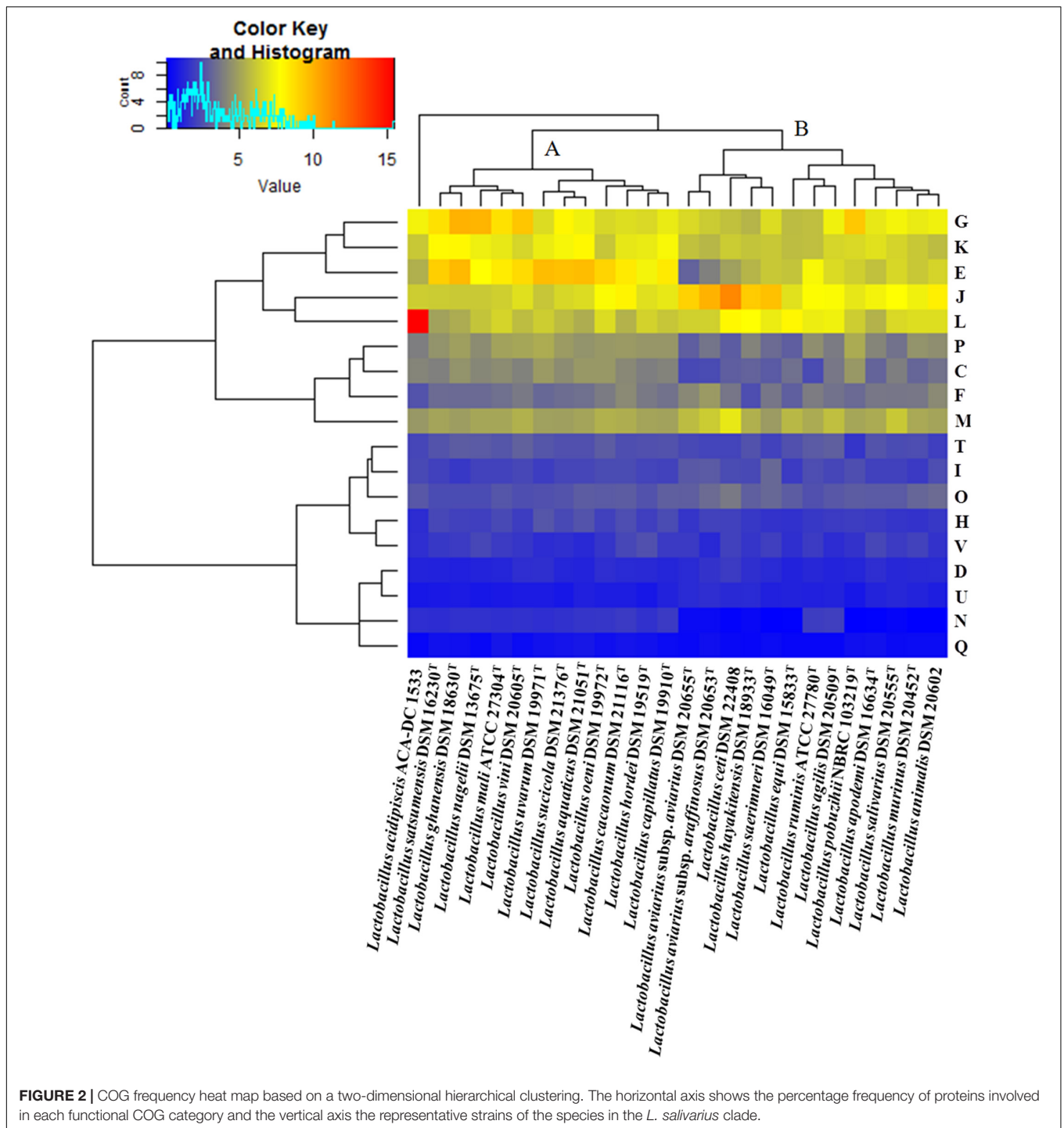


EDGAR (**Supplementary Figure S1**). Results obtained confirmed the clonal relationship among the strains. To evaluate the level of completeness between the *L. acidipiscis* genomes in each of the two pairs of replica strains, we used the genome quality scores from the dBBQs based on the sequence completeness, the tRNA and rRNA score, as well as the number of essential genes predicted in the genome sequence (Wanchai et al., 2017). According to these results, strains KCTC 13900 and JCM 10692^T were found to be more complete than strains DSM 15353 and DSM 15836^T, respectively (**Supplementary Table S2**). For this reason, strains KCTC 13900 and JCM 10692^T were employed for further analysis.

The characteristics of the *L. acidipiscis* ACA-DC 1533 genome were described previously (Asteri et al., 2010; Kazou et al., 2017). The complete chromosomal sequence of the strain was recently re-annotated in RefSeq revealing a total of 2,455 genes including 2,199 protein-coding genes and 172 potential pseudogenes mostly due to frame shifting and internal stop codons (**Figure 3**). Among pseudogenes, hypothetical proteins and mobile elements, such as ISs and transposases, were the most common (**Supplementary Table S3**). The genome also includes six rRNA operons distributed throughout the genome and 63 tRNA genes with the majority located around the five rRNA operons (data not shown).

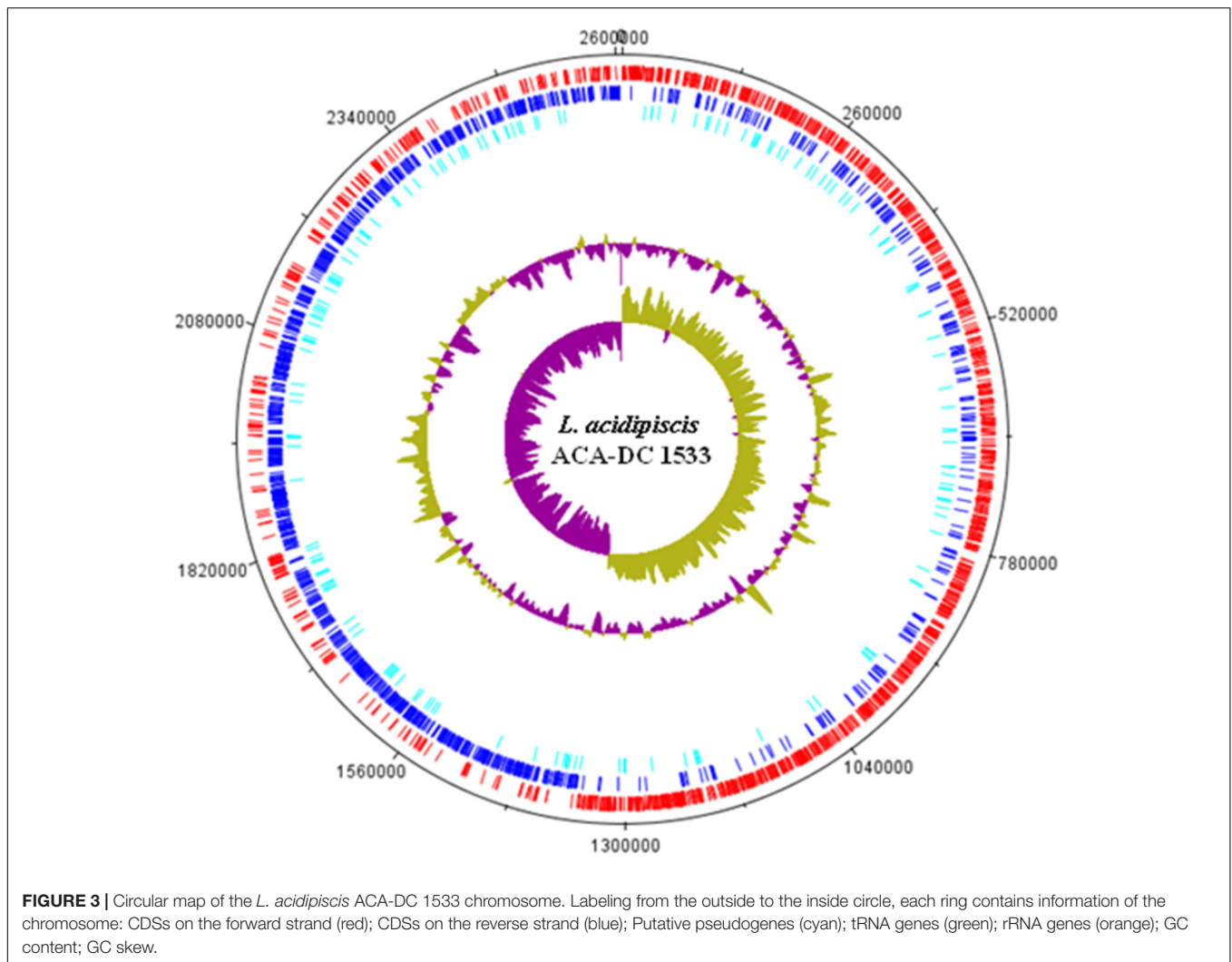
The additional *L. acidipiscis* assemblies of strains JCM 10692^T and KCTC 13900 are fragmented and thus do not allow the determination of their accurate chromosomal size as well as the evaluation of their plasmid content. Nevertheless,

in these assemblies we could detect plasmid sequences after analysis with the r2cat tool, using as templates the three pLAC plasmid sequences (data not shown). Strain ACA-DC 1533 exhibits 2,288 protein-coding genes versus 2,126 and 1,969 for the JCM 10692^T and KCTC 13900 strains, respectively. Analysis with EDGAR revealed that the pan-genome consists of 2,722 genes, with 1,569 and 411 genes belonging to the core- and the dispensable genomes, respectively (**Figure 4A** and **Supplementary Tables S4A,B**). Furthermore, the analysis revealed that singletons represented approximately the 18% of the pan-genome. Strain JCM 10692^T carries the highest number of singletons ($n = 197$) followed by strains ACA-DC 1533 ($n = 157$) and KCTC 13900 ($n = 136$) (**Supplementary Table S4C**). However, such differences may not be readily explained given the differences in completeness among these genomes. We would like to mention that the total number of genes for each strain presented in **Figure 4A** is somewhat lower than the total number of genes annotated for the strain. The missing genes are genes that are not part of the 3-genome or 2-genome cores, but also do not appear in the strictly calculated singletons as they have some second-best BLAST hits, or non-reciprocal-BLAST hits or in general show some similarity to other genes in the dataset that rules them out as singletons as calculated by the EDGAR tool. The distribution of proteins into the COG functional categories is shown in a heat map for the three *L. acidipiscis* strains (**Figure 4B**). Despite their differences in completeness, the three genomes present very similar percentages in each of the COG categories. There was only one exception in replication,



recombination and repair (L) category in which strain ACA-DC 1533 appears to have 15.4% compared to 10.3% and 8.6% for strains JCM 10692^T and KCTC 13900, respectively. As mentioned above, this higher percentage of proteins in the L category for strain ACA-DC 1533 was also evident in the comparison of all species within the *L. salivarius* clade (Figure 2). This difference may again reflect the fragmented nature of *L. acidipiscis* JCM 10692^T and KCTC 13900 genome assemblies.

COG functional classification of the singletons is shown in Figure 5. We could find singletons of the three strains distributed in all COG categories with the majority associated with replication, recombination and repair (L), cell wall/membrane/envelope biogenesis (M), carbohydrate transport and metabolism (G) and transcription (K). The high prevalence of proteins in the L COG category appears again, this time in all three strains, especially strains ACA-DC 1533 and



KCTC 13900. Strain JCM 10692^T appears to have approximately half the singletons in the L COG category, but this may be an artifact deriving from its partial sequence. It is unclear whether genes involved in information storage and processing might have technological implications. It could be suggested though, that the efficiency of central cellular mechanisms like those of the L, M, and K COG categories may provide the strain/species with a competitive advantage in a complex ecosystem. On the contrary, carbohydrate transport and metabolism can have a direct impact on the diversity of ecological niches in which the bacterium can grow.

***In Silico* Evaluation of Motility of *L. acidipiscis* Strains**

From a microbial ecology point of view, motile species may have competitive benefits against non-motile species, regarding e.g., niche colonization and biofilm formation (Neville et al., 2012). Currently, 16 motile *Lactobacillus* species have been recognized in the entire genus, all belonging to the *L. salivarius* clade with the exception of *L. curvatus*, which is a member of the *Lactobacillus*

sakei clade (Cousin et al., 2015). Motility of *L. acidipiscis* has been recently described in strain KCTC 13900 revealing that the 54 proteins involved in flagellum regulation, synthesis, export and chemotaxis are organized in a single operon (Cousin et al., 2015). Annotation of ACA-DC 1533 identified 51 motility genes (LAC1533_RS09635-RS09885) producing a functional flagellar apparatus as also observed by *in vivo* experiments (data not shown). Core-genome analysis revealed that the motility operon is also present in strain JCM 10692^T and flanked by the same genes (Supplementary Table S4B). As shown in Figure 6, alignment of the motility operons of *L. curvatus* NRIC 0822 and *L. acidipiscis* strains KCTC 13900 and ACA-DC 1533 revealed that they are conserved.

GIs Found in *L. acidipiscis* Genomes

HGT is one of the main processes responsible for genome evolution. Genomic fragments acquired by HGT events are characterized as GIs and may have a direct impact on the genome plasticity (Juhás et al., 2009). Here, we focused our analysis in the 13 GIs of the ACA-DC 1533 chromosome identified by

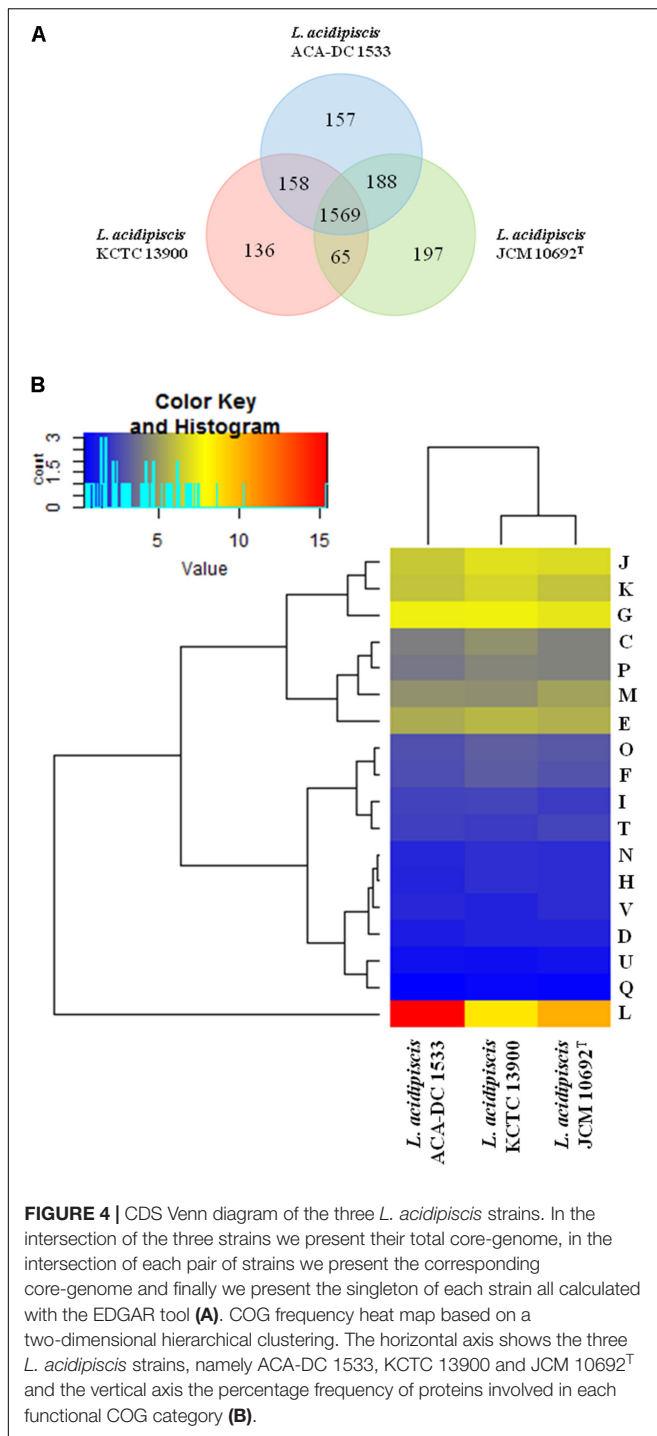


FIGURE 4 | CDS Venn diagram of the three *L. acidipiscis* strains. In the intersection of the three strains we present their total core-genome, in the intersection of each pair of strains we present the corresponding core-genome and finally we present the singleton of each strain all calculated with the EDGAR tool (A). COG frequency heat map based on a two-dimensional hierarchical clustering. The horizontal axis shows the three *L. acidipiscis* strains, namely ACA-DC 1533, KCTC 13900 and JCM 10692^T and the vertical axis the percentage frequency of proteins involved in each functional COG category (B).

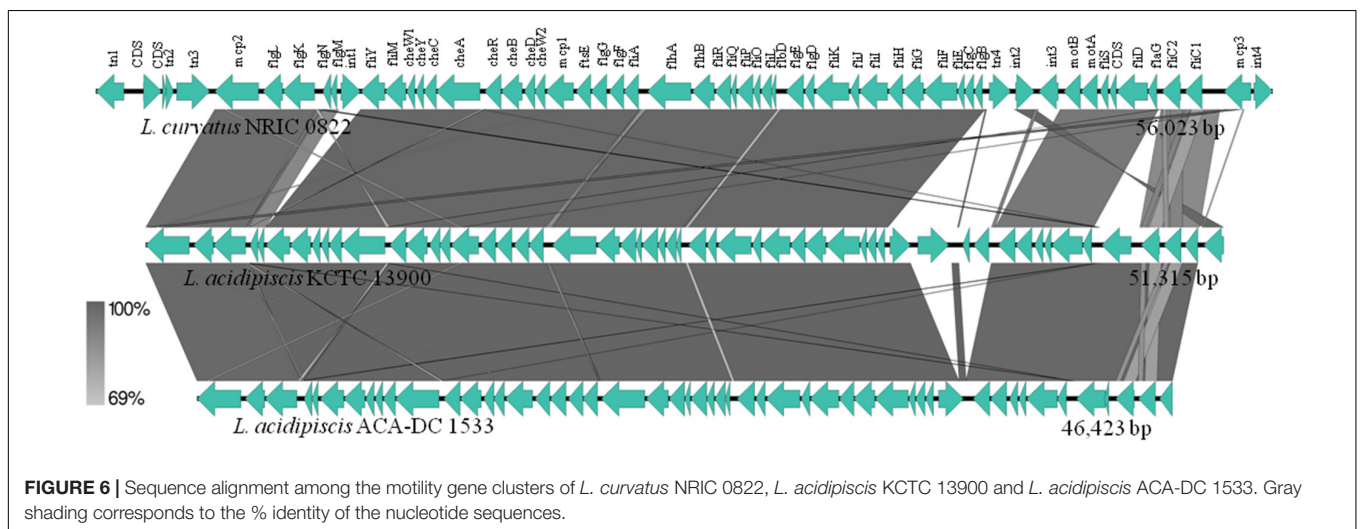
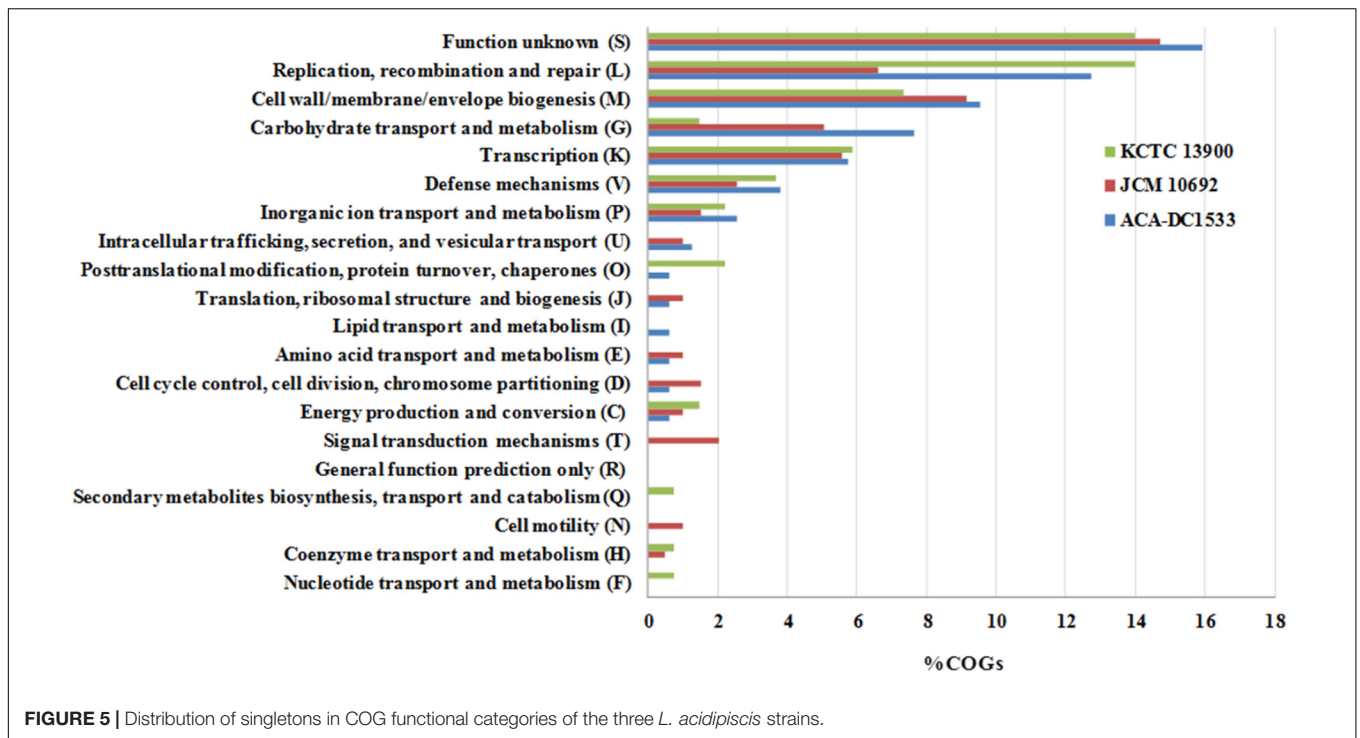
the IslandViewer software tool (Supplementary Figure S2). Of note, GI 9 contains the genome's array of ribosomal proteins (Supplementary Table S5). This is most probably a false positive result, as genes encoding ribosomal proteins have differences in sequence composition compared to regular protein coding genes (Fernández-Gómez et al., 2012) and are thus detected wrongfully by IslandViewer as part of a GI. For this reason, GI 9 was excluded from further analysis. The remaining 12 putative GIs

contain a total of 229 genes and the respective lengths ranging from 4,677 to 36,954 bp. Many of these genes are involved in carbohydrate, lipid and amino acid metabolism as well as in membrane transport systems. According to the pan-genome analysis, GIs 3, 7, and 8 are unique for strain ACA-DC 1533 while GIs 1, 4, and 6 are common in all three *L. acidipiscis* strains, indicating acquisition early in the evolution of the species. It is interesting to note that GI 5 is present in strains ACA-DC 1533 and JCM 10692^T but absent in KCTC 13900. Other GIs are shared among the *L. acidipiscis* strains to a variable degree (Supplementary Table S5).

Prophage Sequences, CRISPR-Cas Systems, RM Systems and TA Systems of *L. acidipiscis* Strains

PHASTER allowed the identification of one intact (1,228,777-1,272,253 bp, from now on called phage 1) and two incomplete prophage regions (1,575,825-1,586,105 bp and 1,802,666-1,830,756 bp) in the ACA-DC 1533 chromosome. Phage 1 contains 53 CDSs, most of which encode hypothetical proteins (approximately 41.5%). Furthermore, phage tail proteins, capsid proteins and *attL/attR* sites flanking the prophage DNA were also identified. Phage 1 is related to several prophages most of which can be found in *Lactobacillus* genomes. Strain JCM 10692^T carries a phage region similar to phage 1 (from now on called phage 2) sharing 30 out of 53 proteins (Supplementary Table S6A). Furthermore, strain KCTC 13900 seems to have an intact prophage region (from now on called phage 3) of 40.8 Kbp length related also to *Lactobacillus* phages (Supplementary Table S6A).

Three CRISPR sequences (i.e., CRISPR1, 2, and 3) were only identified in strain KCTC 13900 (Supplementary Table S6B). BLASTN analysis of all the spacers identified in these three CRISPR-Cas systems showed that several of them, namely spacers 9, 11, 13, 14, 19, 20, and 21 in CRISPR 1 and spacers 5, 14 and 21 in CRISPR 2 had hits in the *Lactobacillus plantarum* virulent phage phiJL-1. Moreover, spacers 22 and 26 in CRISPR 2 had hits in *L. salivarius* plasmids. Since *L. salivarius* strains carrying such plasmids are related to the host environment, this may suggest that *L. acidipiscis* has occupied this niche as well. Most importantly, spacers 1, 3, 5, 6, and 7 in CRISPR 1 and spacer 35 in CRISPR 2 had hits against phage 1 and/or phage 2 genes. Spacers in CRISPRs can reveal aspects of the evolutionary history of their host (Papadimitriou et al., 2014). Thus, it could be hypothesized that strain KCTC 13900 has also been exposed to phage 1 or phage 2 but it was able to acquire immunity through its CRISPR-Cas systems. Our findings may indicate that phages 1 or 2 are abundant in the ecological niches occupied by different *L. acidipiscis* strains or that, despite the different origins of isolation, the three *L. acidipiscis* strains were present in the same ecological niche sometime in the past. Moreover, the presence of phages 1 and 2 in the ACA-DC 1533 and JCM 10692^T genomes, respectively, corroborates with the lack of CRISPR systems in the two strains. However, the presence of prophages in the genomes of *L. acidipiscis* strains may protect them from superinfection by other phages or plasmids (Bondy-Denomy et al., 2016).

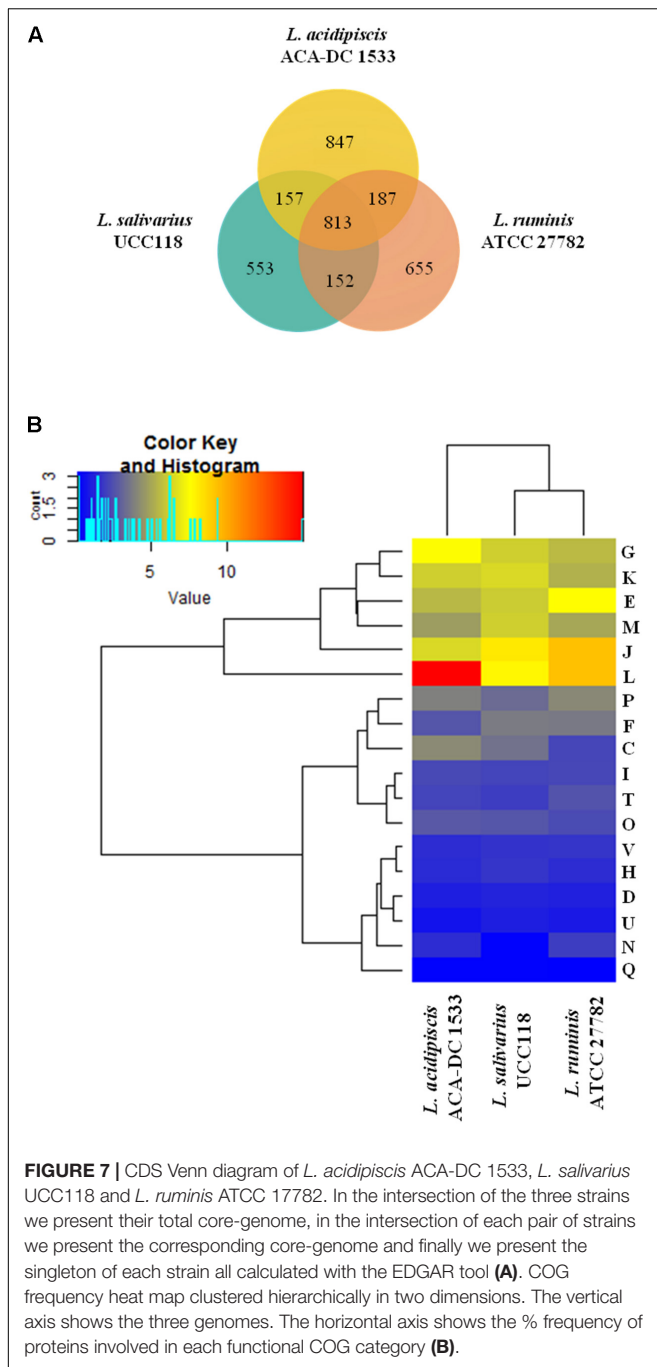


Bacterial defense mechanisms against foreign DNA include RM and TA systems (Darmon and Leach, 2014). Strain ACA-DC 1533 has a type I system that seems to be complete, as it contains the DNA-methyltransferase subunit M (LAC1533_RS04765), the specificity subunits S (LAC1533_RS04770) and R (LAC1533_RS04775), as well as a second type I system (LAC1533_RS01110-RS01130) possibly inactivated, since the restriction subunit R is a potential pseudogene (LAC1533_RS01130). According to the REBASE database, the strain also carries three putative type II RM systems (LAC1533_RS03065, LAC1533_RS05790 and LAC1533_RS08450-RS08455) and two type IV RM systems (LAC1533_RS02780 and LAC1533_RS04790) (Supplementary

Figure S3). Plasmid pLAC3 also carries an *Ava*I RM system. Finally, we looked into TA systems. We concentrated our search on type II TA systems for which TAFinder prediction tool is available. In strain ACA-DC 1533 we found nine TA systems in the chromosome and one in the pLAC2 plasmid (Supplementary Table S7).

Comparative Genomics of *L. acidipiscis* Against *L. salivarius* and *L. ruminis*

To further investigate the lifestyle and/or the technological traits of *L. acidipiscis* ACA-DC 1533, we performed comparative genomic analysis against *L. salivarius* UCC118 and *L. ruminis* ATCC 27782. *L. salivarius* UCC118 was chosen as the

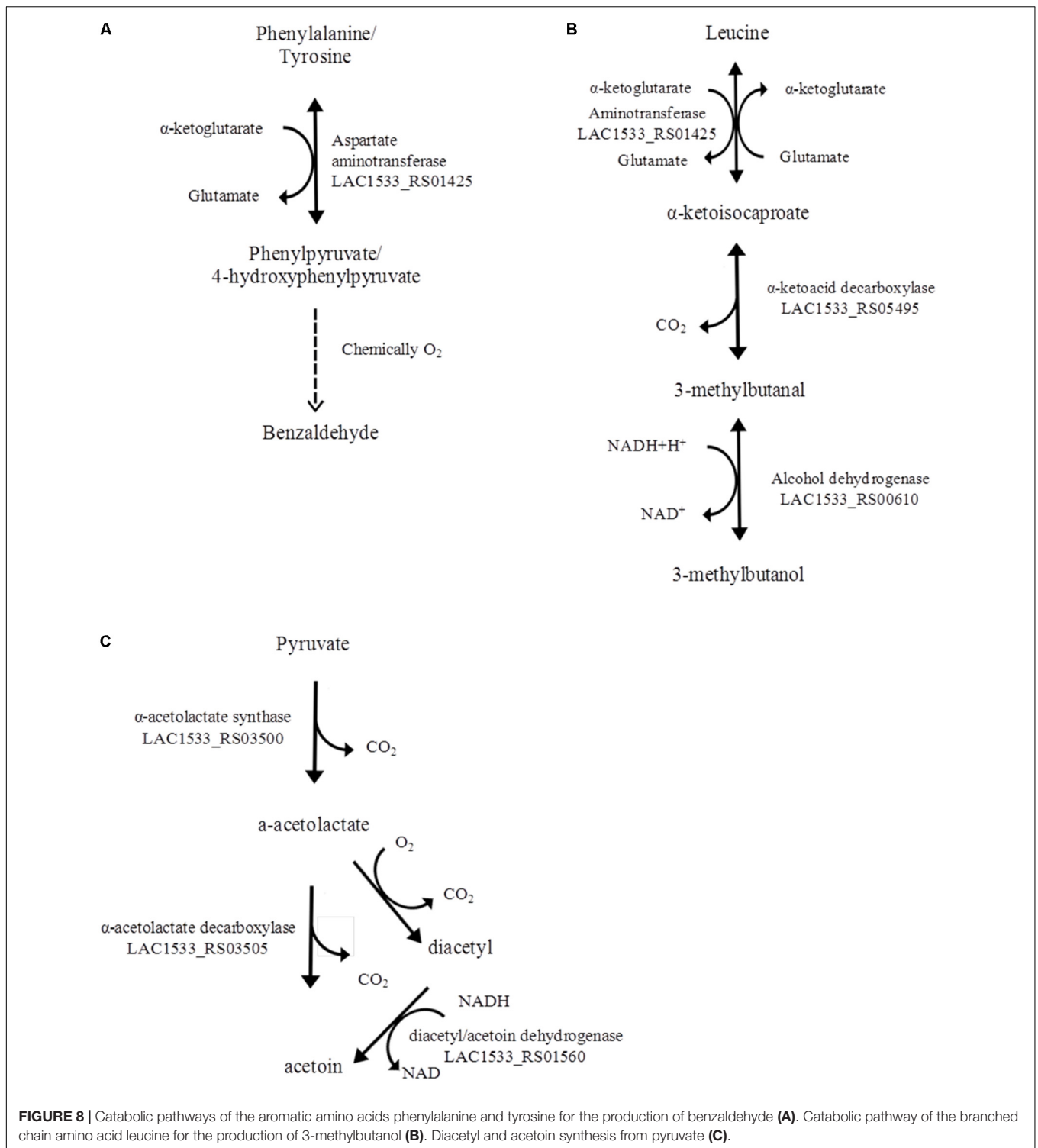


representative strain of the species since it is the first sequenced and presumably the best characterized strain of the clade (Harris et al., 2017). The comparison was performed initially at the chromosome level since the chromosomes of all three strains are completely sequenced. *L. salivarius* UCC118 was isolated from the human ileal-caecal region and comprises a chromosome of 1.8 Mbp and three plasmids, one of which is a megaplasmid of 242 Kbp (Claesson et al., 2006). *L. ruminis* ATCC 27782 isolated from the bovine rumen has a chromosome size of 2.1 Mbp with no plasmids (Forde et al., 2011). As mentioned above,

L. acidipiscis ACA-DC 1533 has a chromosome of 2.6 Mbp, which is the largest among the three species. *L. acidipiscis* ACA-DC 1533 and *L. ruminis* ATCC 27782 exhibited the highest number of potential pseudogenes, i.e., 7.3 and 9.0%, respectively in contrast to the 2.8% of *L. salivarius* UCC188. However, other complete *L. salivarius* chromosomes exhibit a variable percentage of potential pseudogenes, up to 6.6% (Supplementary Table S8). Taking this observation into account, it seems that pseudogenes may not be constant among strains of the same species and thus the existence of only one complete chromosomal sequence for *L. acidipiscis* and *L. ruminis* are not enough to comment about their overall genome decay at the species level. Nevertheless, *L. acidipiscis* ACA-DC 1533 and *L. ruminis* ATCC 27782 appear to have undergone genome decay to an extent that is relatively restricted, at least when compared to the genome decay of highly specialized dairy lactobacilli like *L. delbrueckii* subsp. *bulgaricus* (van de Guchte et al., 2006).

Our analysis also revealed that the number of common proteins among the three species is 813, higher than that calculated for the entire *L. salivarius* clade as analyzed above (Figure 7A and Supplementary Table S9A). *L. acidipiscis* ACA-DC 1533 seems to carry the highest number of unique genes ($n = 847$) mostly encoding hypothetical proteins, transposases, ABC transporters, PEP-PTS and membrane transport proteins (Supplementary Table S9B). Similarly to Figure 4A, the total number of genes for each strain presented is somewhat lower than the total number of genes annotated for the strain since some genes cannot be assigned neither in the singletons nor in the 3-genome or 2-genome cores for the reason presented above. Furthermore, there is no extensive synteny among the three species as observed during full-length chromosome alignments created by progressiveMAUVE (Supplementary Figure S4). The analysis revealed a high number of LCBs with a quite short average length. Several studies based on comparative genomics among *Lactobacillus* species have established the genomic diversity of the *Lactobacillus* genus, which is higher compared to that of a typical bacterial family (Sun Z. et al., 2015; Martino et al., 2016).

The distribution of proteins into the COG functional categories for the three species is shown in Figure 7B. As expected, *L. acidipiscis* ACA-DC 1533 chromosome contained more proteins compared to *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 in the L COG category owing to an inflated number of transposases and reverse transcriptases. Inspection of each of the two categories of gene products revealed that they may contain in some instances identical paralogs, but this is not always the case. The biological reason behind this observation is not clear. However, considering that both *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 chromosomes are completely sequenced and that both *L. acidipiscis* ACA-DC 1533 and *L. ruminis* ATCC 27782 are annotated with the same pipeline in RefSeq, the possibility that this difference is some type of artifact is rather unlikely. Another obvious difference was the absence of proteins in the cell motility (N) COG category from the *L. salivarius* UCC118 chromosome. In all other COG categories,



the distribution of proteins was at a comparable level among the three strains.

We also compared plasmid sequences of *L. acidipiscis* ACA-DC 1533 and *L. salivarius* UCC118. It has been shown for the latter that important housekeeping genes may be carried in its plasmids (Harris et al., 2017). In the case of *L. acidipiscis*

plasmids most of the proteins were hypothetical. Nevertheless, we were able to identify some genes encoding proteins that may be important for the physiology, metabolism and/or the technological properties of the strain. For example, we determined the presence of carbohydrate and ion transporters (**Supplementary Table S10**), putative carbohydrate metabolizing

enzymes (please see below), and an arsenate reductase. In addition and as mentioned above, plasmids of *L. acidipiscis* ACA-DC 1533 carry an *Ava*I RM system and a type II TA system (**Supplementary Table S7**).

Glycobiome Analysis of *L. acidipiscis*, *L. salivarius*, and *L. ruminis*

The glycobiomes of *L. acidipiscis* ACA-DC 1533, *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 were investigated using dbCAN. According to the analysis, *L. acidipiscis* ACA-DC 1533 had the largest glycobiome with 85 enzymes involved in carbohydrate metabolism, followed by *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 with 78 and 68 enzymes, respectively (**Supplementary Table S11**). Among the 85 enzymes, 37 were identified as GHs, 21 as GTs, 13 as CEs and 14 as CBMs. Compared to the 37 GHs of *L. acidipiscis* ACA-DC 1533, *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 contained 27 and 26 GHs, respectively. Among the GH families identified in the *L. acidipiscis* ACA-DC 1533, *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 genomes, GH 13 was the most pronounced containing mainly enzymes with plant substrate specificity (Crost et al., 2013). Indeed, the carbohydrate fermentation profile of *L. acidipiscis* ACA-DC 1533 using the API 50 CHL stripes (**Supplementary Table S12**) and *L. salivarius* UCC118 (Li et al., 2006) showed that the two strains were able to ferment a number of carbohydrates of plant origin, i.e., L-arabinose, D-ribose, D-cellobiose, D-trehalose, D-glucose, D-fructose, D-mannitol, D-sorbitol, and D-saccharose. Furthermore, several GH families, namely GH 35, GH 38, GH 46, GH 70, and GH 76, were unique for the *L. acidipiscis* ACA-DC 1533 genome indicating that the bacterium presumably requires these enzymes in its ecological niche, which might be different to that of *L. salivarius* UCC118 and *L. ruminis* ATCC 27782. Interestingly, the presence of a beta-galactosidase (GH 35) and two 6-phospho-beta-galactosidase genes (GH 1) in the *L. acidipiscis* ACA-DC 1533 genome could probably be required for growth in milk. *L. acidipiscis* ACA-DC 1533 genome seems to contain also the highest number of CBM modules in family 50 compared to the *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 genomes. CBM 50 modules are commonly found in bacterial lysins having a peptidoglycan binding function and a contribution to cell division (Visweswaran et al., 2013). Similarly to what has been reported previously for *L. salivarius* UCC118 (Harris et al., 2017) and according to our analysis, part of the glycobiome of both *L. salivarius* and *L. acidipiscis* ACA-DC 1533 resides in their plasmids. Specifically for *L. acidipiscis*, we found two GT 4 in plasmid pLAC2. It seems plausible to state that diversity of the plasmid glycobiome in strains of *L. salivarius* is significantly more rich than that of *L. acidipiscis* perhaps due to the presence of the megaplasmid. Moreover, analysis using the TransportDB database identified 47 potential sugar specific PTS transport proteins in the *L. acidipiscis* ACA-DC 1533 genome (3 on pLAC2) and 25 and 16 potential PTS transport proteins for *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 genomes, respectively (**Supplementary Table S10**).

Proteolytic System of *L. acidipiscis*, *L. salivarius*, and *L. ruminis*

The proteolytic system of lactic acid bacteria consists of cell-wall bound proteinases, which initiate the degradation of caseins, peptide and amino acid transport systems and a pool of intracellular peptidases, which further degrade the peptides to shorter peptides and free amino acids (Liu et al., 2010). The proteolytic system of the three *L. acidipiscis* strains, *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 was investigated according to the scheme of Liu and co-workers (Liu et al., 2010) (**Supplementary Table S13**). The cell-wall bound proteinase (PrpP), the aminopeptidase A (PepA), the endopeptidases PepE/PepG and the proline peptidase PepL were missing from all strains. It is worth mentioning that PrpP gene is intact in plasmid pR1 of *L. salivarius* strain Ren (Sun E. et al., 2015). The rest of the peptidases were found in up to three copies per genome. Furthermore, *L. acidipiscis* ACA-DC 1533 and *L. ruminis* ATCC 27782 carried one oligopeptide ABC transport system (Opp), which was missing from the *L. salivarius* UCC118 genome. Interestingly, the Opp operon is present in *L. acidipiscis* ACA-DC 1533 and JCM 10692^T but absent in KCTC 13900. On the contrary, a di/tripeptide ABC transport system (Dpp) and a DtpT transporter of di- and tri-peptides were found in the three species (including all three *L. acidipiscis* strains). However, it is worth noting that the DppD protein of *L. acidipiscis* KCTC 13900 is a potential pseudogene inactivating the entire Dpp system which deserves further investigation. Moreover, *L. acidipiscis* ACA-DC 1533 chromosome seems to contain 17 amino acid ABC transport proteins, while *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 chromosomes only 11 and 10, respectively. Even though the five *Lactobacillus* chromosomes and/or genomes carry a number of peptide and amino acid transporters as well as several intracellular peptidases, the absence of PrpP indicates that the strains may not directly hydrolyze large protein molecules, but they may take advantage of peptides and free amino acids already available in their ecological niche.

Miscellaneous Genomic Features Deriving From the Comparison Among *L. acidipiscis* ACA-DC 1533, *L. salivarius* UCC118, and *L. ruminis* ATCC 27782

We also focused our analysis to IS elements that may contribute in bacterial genome evolution, to transport proteins which allow the transport of the substances in and out of the cell, as well as to RPs that control gene expression.

IS elements of *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 have been previously identified (Claesson et al., 2006) but we have updated the analysis using the latest version of ISSaga and the most recent annotation files for the two strains. In the chromosomes of *L. acidipiscis* ACA-DC 1533, *L. salivarius* UCC118 and *L. ruminis* ATCC 27782, a total of 53, 10 and 30 IS elements were predicted with ISSaga, respectively (**Supplementary Table S14**). The higher number of IS elements in the chromosome of *L. acidipiscis* ACA-DC 1533 may suggest a higher potential for genome plasticity compared

to the *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 chromosome. The majority of IS elements in the *L. acidipiscis* ACA-DC 1533 chromosome belong to the ISL3 and IS982 families which were also previously identified in food related lactobacilli like *Lactobacillus delbrueckii* subsp. *bulgaricus* and *Lactobacillus helveticus*, respectively (Germond et al., 1995; Callanan et al., 2005).

Furthermore, the *L. acidipiscis* ACA-DC 1533 genome contains 287 transport proteins compared to 240 and 238 of *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 genomes, respectively. They mainly belong to the ABC superfamily and to the MFS (**Supplementary Table S10**). Additional analysis of the *L. acidipiscis* ACA-DC 1533 genome revealed 17 potential glycine/betaine transport proteins organized in at least five distinct genomic loci. The glycine/betaine transport system may be necessary to overcome osmotic stress since *L. acidipiscis* is a salt-tolerant species owning strains able to grow in the presence of even 12% NaCl (our unpublished results; Tanasupawat et al., 2000; Romeo et al., 2003; Pot et al., 2014).

RPs include TCSs and TFs. TCSs are the most abundant phosphorylation-dependent signal transduction systems in prokaryotes and typically comprise a membrane-bound HK and a RR (Barakat et al., 2013). On the other hand, TFs contain TRs, OCSs, RRs and SFs. Analysis of *L. acidipiscis* ACA-DC 1533 and *L. salivarius* UCC118 identified six HKs and seven RRs for both strains. Analysis of *L. ruminis* ATCC 27782 chromosome revealed seven HKs and 10 RRs. Furthermore, the *L. acidipiscis* ACA-DC 1533 chromosome contained the highest number of TFs among the three strains analyzed, including 68 TRs, 28 OCSs, five RRs, six SFs and 19 ODPs, most of which were unclassified (**Supplementary Table S15**). The higher number of TFs in the *L. acidipiscis* compared to the other two species may suggest a more intricate regulation of gene expression and perhaps an increased interaction with the environment.

Assessing the Probiotic and Technological Properties of *L. acidipiscis* ACA-DC 1533

Initially, we investigated the probiotic potential of *L. acidipiscis* ACA-DC 1533 based on the available information for *L. salivarius* UCC118 which has been extensively studied as a probiotic strain (Neville and O'Toole, 2010). The *L. salivarius* UCC118 genome contains a bile-salt hydrolase (Claesson et al., 2006) and two EPS clusters associated with the strain's probiotic activity (Harris et al., 2017). These traits were absent from the *L. acidipiscis* ACA-DC 1533 genome. In addition, proteins that may play a role in the interaction of *L. salivarius* UCC118 with the host, may include mucus-, collagen-, salivary agglutinin- and epithelial-binding proteins, as well as enterococcal surface proteins (van Pijkeren et al., 2006; O'Shea et al., 2012). All these proteins are sortase-dependent surface proteins which were either absent from the *L. acidipiscis* ACA-DC 1533 genome or were characterized as potential pseudogenes. The only exception identified was a fibrinogen/fibronectin-binding protein, similar to that of *L. salivarius* UCC118 (Collins et al., 2012) that

was also present in the *L. acidipiscis* ACA-DC 1533 genome. Furthermore, analysis of the *L. acidipiscis* ACA-DC 1533 genome with the BAGEL tool did not predict any bacteriocin gene, in contrast to the *L. salivarius* UCC118 genome, which produces the two-component class II bacteriocin Abp118 (Flynn et al., 2002). BAGEL also predicted in *L. acidipiscis* JCM 10692^T three potential structural genes coding for pediocin, sakacin P and carnocin like bacteriocins (the last being a potential pseudogene) and some accessory genes (e.g., immunity, transfer, and maturation) and further experimental testing for their production needs to be performed.

We then investigated aspects of the technological potential of *L. acidipiscis* ACA-DC 1533 taking into account that Asteri and co-workers showed that the major volatile/flavor metabolites produced by this strain when grown in RSM and MRS, were 3-methylbutanal, 3-methylbutanol, benzaldehyde and acetoin (Asteri et al., 2009). The majority of the aforementioned metabolites produced by *L. acidipiscis* ACA-DC 1533 are degradation products of amino acids (**Figure 8**). In particular, benzaldehyde can be formed from two aromatic amino acids, namely phenylalanine and tyrosine, using an enzymatic and a non-enzymatic step (Nierop Groot and de Bont, 1998; Fernandez and Zuniga, 2006). Moreover, 3-methylbutanal and 3-methylbutanol are catabolic products of the branched-chain amino acid leucine (Fernandez and Zuniga, 2006). The α -ketoacid decarboxylase and the alcohol dehydrogenase involved in the leucine catabolism pathway were found to be present in the three *L. acidipiscis* genomes but absent from *L. salivarius* UCC118 and *L. ruminis* ATCC 27782. On the contrary, aspartate aminotransferase, which catalyzes the transamination of phenylalanine and tyrosine, was present in all the *Lactobacillus* genomes analyzed. Many studies have been shown that the amino acid degradation products, especially those deriving from the branched-chain, aromatic and sulfur-containing amino acids, are regarded as significant flavor compounds in several cheese varieties (Ardö, 2006; Liu et al., 2008; Afzal et al., 2017). Furthermore, acetoin, which was produced by *L. acidipiscis* ACA-DC 1533, can be formed from pyruvate using two alternative pathways. Pyruvate, which derives from glycolysis, is converted into α -acetolactate by α -acetolactate synthase (LAC1533_RS03500). α -Acetolactate is then catabolized either to acetoin by α -acetolactate decarboxylase (LAC1533_RS03505) or to diacetyl in the presence of oxygen. Finally, diacetyl/acetoin dehydrogenase (LAC1533_RS01560) catalyzes the conversion of diacetyl to acetoin (Celinska and Grajek, 2009). It should be mentioned that diacetyl was not detected as a volatile metabolite of *L. acidipiscis* ACA-DC 1533 in the work of Asteri et al. (2009). However, the presence of diacetyl/acetoin dehydrogenase in the ACA-DC 1533 genome could probably mean that by the time of sampling diacetyl was fully converted into acetoin. Given that *L. acidipiscis* ACA-DC 1533, along with *L. rennini*, were the only species found in Kopanisti cheese, the production of the above mentioned metabolites by *L. acidipiscis* ACA-DC 1533 via amino acid catabolism may contribute to the characteristic piquant flavor of Kopanisti cheese (Yvon and Rijnen, 2001; Asteri et al., 2009; Donnelly, 2016).

CONCLUSION

The *L. salivarius* clade includes mainly commensal species and it has been suggested that several strains may have probiotic properties (Neville and O'Toole, 2010). In this study, we analyzed the available genomes of *L. acidipiscis*, a species within the *L. salivarius* clade that until today has mainly been isolated from fermented foods of dairy or other origin. We wanted to examine whether *L. acidipiscis* is also a commensal that is transferred to the ecosystem of fermented foods accidentally from the host. Furthermore, we wanted to investigate the probiotic and/or the technological potential of the species. We sequenced the genome of strain ACA-DC 1533, originally isolated from Kopanisti, a traditional spread-type cheese that is highly salted and particularly piquant (Kazou et al., 2017). Our investigation suggested that *L. acidipiscis* has a relatively large genome compared to other species of the *L. salivarius* clade (at least those of *L. salivarius* and *L. ruminis*) with a relatively restricted percentage of pseudogenes. These findings along with the observation that *L. acidipiscis* possesses a high number of glycozyme enzymes may indicate an ability to occupy versatile environments and has not been evolved toward a specific ecological niche. Perhaps, adaptation to a nutrient-rich niche would have been more consistent with a smaller genome size, more typically like that of *L. salivarius*. Interestingly, *L. acidipiscis* strains ACA-DC 1533 and JCM 10692^T appear to be more related compared to strain KCTC 13900 based on the presence/absence distribution of a number of genetic traits like prophages and CRISPRs. *L. acidipiscis* ACA-DC 1533 does not seem to present any evident probiotic trait at the genomic level. Besides the absence of several genes that have been related to probiotic properties of *L. salivarius* UCC118 preliminary experiments with strain *L. acidipiscis* ACA-DC 1533 on human PBMCs did not reveal an increased interleucine-10/interleucine-12 ratio, indicative of a potential to stimulate a Treg response (our unpublished results). Some probiotic properties have been suggested for specific *L. acidipiscis* strains like an antiproliferative effect against Caco-2 cells (Thamacharoensuk et al., 2017) or the improvement of feed conversion efficiency in broiler chickens (Altaher et al., 2015). Thus, further *in silico* and/or experimental assessment of the probiotic properties of *L. acidipiscis* may be required. In addition, *L. acidipiscis* is able to grow in the dairy environment since it can ferment lactose, it possesses a complete proteolytic system for the degradation of milk proteins (without carrying a cell-envelope proteinase) and it can produce volatile compounds during the catabolism of amino acids that may contribute to the flavor of the final product. Intriguingly, *L. acidipiscis* is also considered a spoiler in vinegar and soy sauce. For this reason, technological steps to prohibit its growth in certain fermented foods need to be devised. Further research is needed with different species of the *L. salivarius* clade like the newly sequenced *Lactobacillus agilis* to better appreciate the mechanisms underlining the adaptation to the host and/or the food environment. Sequencing of more strains/species would provide invaluable information about the ecology of this important clade within the *Lactobacillus* genus.

AUTHOR CONTRIBUTIONS

MK and VA performed genome analysis and participated in the writing of the manuscript. JB and BP performed genome analysis. ET conceived the project and participated in the writing of the manuscript. KP conceived the project, performed genome analysis, and participated in the writing of the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was co-financed by the European Social Fund and the National resources EPEAEK and YPEPTH through the Thales project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2018.01244/full#supplementary-material>

FIGURE S1 | Heat map of ANI values among the five sequenced *L. acidipiscis* strains.

FIGURE S2 | Circular map of the *L. acidipiscis* ACA-DC 1533 chromosome as generated by IslandViewer 4. Highlighted regions correspond to GIs. GIs are colored within the circular map according to the prediction method used: GIs in orange were predicted with SIGI-HMM, GIs in blue with IslandPath-DIMOB and the integrated GIs are presented on the periphery in red. The black line plot corresponds to the GC content (%) of the chromosomal sequence.

FIGURE S3 | Circular map of the *L. acidipiscis* ACA-DC 1533 chromosome highlighting the predicted RM systems by the REBASE database. The symbols are color coded as indicated at the bottom of the figure.

FIGURE S4 | Chromosome alignments among the *L. acidipiscis* ACA-DC 1533, *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 strains generated by progressiveMAUVE. Locally collinear blocks (LCBs) of conserved sequences are presented by the same color (white corresponds to the strain-specific regions).

TABLE S1 | General information of the strains analyzed in this study.

TABLE S2 | Level of completeness among the four *L. acidipiscis* strains as calculated by dBBQs.

TABLE S3 | Pseudogenes identified in the *L. acidipiscis* ACA-DC 1533 chromosome.

TABLE S4 | (A) Pan-genome analysis among the three *L. acidipiscis* strains calculated with the EDGAR software. (B) Core-genome among the three *L. acidipiscis* strains calculated with the EDGAR software. (C) Singletons of the three *L. acidipiscis* strains calculated with the EDGAR software.

TABLE S5 | Genes within genomic islands of *L. acidipiscis* ACA-DC 1533. Common genes within the rest of *L. acidipiscis* strains were predicted from the pan-genome analysis with the EDGAR software.

TABLE S6 | (A) Prophage regions identified in the three *L. acidipiscis* strains using the PHASTER software. (B) CRISPR systems identified in the *L. acidipiscis* KCTC 13900 genome using the CRISPRFinder tool.

TABLE S7 | TA systems predicted in the chromosome and pLAC2 plasmid of *L. acidipiscis* ACA-DC 1533 genome.

TABLE S8 | Percentage of potential pseudogenes identified in the chromosomes of the completed sequenced genomes of *L. salivarius* clade.

TABLE S9 | (A) Core-genome among the *L. acidipiscis* ACA-DC 1533, *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 chromosomes as calculated with the EDGAR software. **(B)** Singletons among the *L. acidipiscis* ACA-DC 1533, *L. salivarius* UCC118 and *L. ruminis* ATCC 27782 chromosomes as calculated with the EDGAR software.

TABLE S10 | Transporters of *L. acidipiscis* ACA-DC 1533 chromosome and pLAC2 plasmid identified by TransportDB.

TABLE S11 | Putative enzymes involved in carbohydrate metabolism of *L. acidipiscis* ACA-DC 1533, *L. salivarius* UCC118, and *L. ruminis* ATCC 27782 chromosomes.

TABLE S12 | Acid production by *L. acidipiscis* ACA-DC 1533 using API 50CHL stripes.

TABLE S13 | The proteolytic system of *L. acidipiscis*, *L. salivarius*, and *L. ruminis* chromosomes and/or genomes.

TABLE S14 | Insertion sequences (ISs) identified among the *L. acidipiscis* ACA-DC 1533, *L. salivarius* UCC118, and *L. ruminis* ATCC 27782 chromosomes using the ISSaga platform.

TABLE S15 | Regulatory proteins identified in the chromosomes of *L. acidipiscis* ACA-DC 1533, *L. salivarius* UCC118, and *L. ruminis* ATCC 27782 using P2RP web server.

REFERENCES

- Afzal, M. I., Ariceaga, C. C., Boulahya, K.-A., Jacquot, M., Delaunay, S., and Cailliez-Grimal, C. (2017). Biosynthesis and role of 3-methylbutanal in cheese by lactic acid bacteria: major metabolic pathways, enzymes involved, and strategies for control. *Crit. Rev. Food Sci. Nutr.* 57, 399–406. doi: 10.1080/10408398.2014.893502
- Altaher, Y., Jahromi, M., Ebrahim, R., Zulkifli, I., and Liang, J. (2015). *Lactobacillus pentosus* Ita23 and *L. acidipiscis* Ita44 enhance feed conversion efficiency and beneficial gut microbiota in broiler chickens. *Rev. Bras. Cienc. Avic.* 17, 159–164. doi: 10.1590/1516-635x1702159-164
- An, C., Takahashi, H., Kimura, B., and Kuda, T. (2010). Comparison of PCR-DGGE and PCR-SSCP analysis for bacterial flora of Japanese traditional fermented fish products, *aji-narezushi* and *iwashi-nukazuke*. *J. Sci. Food Agric.* 90, 1796–1801. doi: 10.1002/jsfa.4015
- Arasu, M. V., Al-Dhabi, N. A., Rejiniemon, T. S., Lee, K. D., Huxley, V. A. J., Kim, D. H., et al. (2015). Identification and characterization of *Lactobacillus brevis* P68 with antifungal, antioxidant and probiotic functional properties. *Indian J. Microbiol.* 55, 19–28. doi: 10.1007/s12088-014-0495-3
- Ardö, Y. (2006). Flavour formation by amino acid catabolism. *Biotechnol. Adv.* 24, 238–242. doi: 10.1016/j.biotechadv.2005.11.005
- Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21. doi: 10.1093/nar/gkw387
- Asteri, I.-A., Kittaki, N., and Tsakalidou, E. (2010). The effect of wild lactic acid bacteria on the production of goat's milk soft cheese. *Int. J. Dairy Technol.* 63, 234–242. doi: 10.1111/j.1471-0307.2010.00564.x
- Asteri, I.-A., Robertson, N., Kagli, D.-M., Andrewes, P., Nychas, G., Coolbear, T., et al. (2009). Technological and flavour potential of cultures isolated from traditional Greek cheeses – A pool of novel species and starters. *Int. Dairy J.* 19, 595–604. doi: 10.1016/j.jidairy.2009.04.006
- Barakat, M., Ortet, P., and Whitworth, D. E. (2013). P2RP: a Web-based framework for the identification and analysis of regulatory proteins in prokaryotic genomes. *BMC Genomics* 14:269. doi: 10.1186/1471-2164-14-269
- Bernardeau, M., Vernoux, J. P., Henri-Dubernet, S., and Gueguen, M. (2008). Safety assessment of dairy microorganisms: the *Lactobacillus* genus. *Int. J. Food Microbiol.* 126, 278–285. doi: 10.1016/j.ijfoodmicro.2007.08.015
- Bertelli, C., Laird, M. R., Williams, K. P., Lau, B. Y., Hoard, G., Winsor, G. L., et al. (2017). IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* 45, W30–W35. doi: 10.1093/nar/gkx343
- Blom, J., Albaum, S. P., Doppmeier, D., Puhler, A., Vorholter, F. J., Zakrzewski, M., et al. (2009). EDGAR: a software framework for the comparative analysis of prokaryotic genomes. *BMC Bioinformatics* 10:154. doi: 10.1186/1471-2105-10-154
- Bondy-Denomy, J., Qian, J., Westra, E. R., Buckling, A., Guttman, D. S., Davidson, A. R., et al. (2016). Prophages mediate defense against phage infection through diverse mechanisms. *ISME J.* 10, 2854–2866. doi: 10.1038/ismej.2016.79
- Callanan, M. J., Beresford, T. P., and Ross, R. P. (2005). Genetic diversity in the Lactose operons of *Lactobacillus helveticus* strains and its relationship to the role of these strains as commercial starter cultures. *Appl. Environ. Microbiol.* 71, 1655–1658. doi: 10.1128/AEM.71.3.1655-1658.2005
- Cannon, J. P., Lee, T. A., Bolanos, J. T., and Danziger, L. H. (2005). Pathogenic relevance of *Lactobacillus*: a retrospective review of over 200 cases. *Eur. J. Clin. Microbiol. Infect. Dis.* 24, 31–40. doi: 10.1007/s10096-004-1253-y
- Cao, J., Yang, J., Hou, Q., Xu, H., Zheng, Y., Zhang, H., et al. (2017). Assessment of bacterial profiles in aged, home-made Sichuan paocai brine with varying titratable acidity by PacBio SMRT sequencing technology. *Food Control* 78(Suppl. C), 14–23. doi: 10.1016/j.foodcont.2017.02.006
- Carver, T., Thomson, N., Bleasby, A., Berriman, M., and Parkhill, J. (2009). DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics* 25, 119–120. doi: 10.1093/bioinformatics/btn578
- Celinska, E., and Grajek, W. (2009). Biotechnological production of 2,3-butanediol-current state and prospects. *Biotechnol. Adv.* 27, 715–725. doi: 10.1016/j.biotechadv.2009.05.002
- Cheng, L., Lin, W., Li, P., Huang, J., and Luo, L. (2014). Comparison of microbial communities between normal and swollen canned soy sauces using nested PCR-denaturing gradient gel electrophoresis, HPLC and plate techniques. *Int. J. Food Sci. Technol.* 49, 2499–2505. doi: 10.1111/ijfs.12575
- Chenoll, E., Codoñer, F. M., Martínez-Blanch, J. F., Acevedo-Piérart, M., Ormeño, M. L., Ramón, D., et al. (2016). Complete genome sequence of the probiotic strain *Lactobacillus salivarius* LPM01. *Genome Announc.* 4:e01319-16. doi: 10.1128/genomeA.01319-16
- Claesson, M. J., Li, Y., Leahy, S., Canchaya, C., van Pijkeren, J. P., Cerdeno-Tarraga, A. M., et al. (2006). Multireplicon genome architecture of *Lactobacillus salivarius*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6718–6723. doi: 10.1073/pnas.0511060103
- Collins, J., van Pijkeren, J.-P., Svensson, L., Claesson, M. J., Sturme, M., Li, Y., et al. (2012). Fibrinogen-binding and platelet-aggregation activities of a *Lactobacillus salivarius* septicaemia isolate are mediated by a novel fibrinogen-binding protein. *Mol. Microbiol.* 85, 862–877. doi: 10.1111/j.1365-2958.2012.08148.x
- Cousin, F. J., Lynch, S. M., Harris, H. M., McCann, A., Lynch, D. B., Neville, B. A., et al. (2015). Detection and genomic characterization of motility in *Lactobacillus curvatus*: confirmation of motility in a species outside the *Lactobacillus salivarius* clade. *Appl. Environ. Microbiol.* 81, 1297–1308. doi: 10.1128/AEM.03594-14
- Crost, E. H., Tailford, L. E., Le Gall, G., Fons, M., Henrissat, B., and Juge, N. (2013). Utilisation of mucin glycans by the human gut symbiont *Ruminococcus gnavus* is strain-dependent. *PLoS One* 8:e76341. doi: 10.1371/journal.pone.0076341
- Darling, A. E., Mau, B., and Perna, N. T. (2010). ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. doi: 10.1371/journal.pone.0011147
- Darmon, E., and Leach, D. R. (2014). Bacterial genome instability. *Microbiol. Mol. Biol. Rev.* 78, 1–39. doi: 10.1128/mbr.00035-13
- Donnelly, C. (2016). *The Oxford Companion to Cheese*. Oxford: Oxford University Press.
- Duar, R. M., Lin, X. B., Zheng, J., Martino, M. E., Grenier, T., Perez-Munoz, M. E., et al. (2017). Lifestyles in transition: evolution and natural history of the genus *Lactobacillus*. *FEMS Microbiol. Rev.* 41(Suppl._1), S27–S48. doi: 10.1093/femsre/fux030
- Elbourne, L. D. H., Tetu, S. G., Hassan, K. A., and Paulsen, I. T. (2017). TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Res.* 45, D320–D324. doi: 10.1093/nar/gkw1068
- Fernandez, M., and Zuniga, M. (2006). Amino acid catabolic pathways of lactic acid bacteria. *Crit. Rev. Microbiol.* 32, 155–183. doi: 10.1080/10408410600880643
- Fernández-Gómez, B., Fernández-Guerra, A., Casamayor, E. O., González, J. M., Pedrós-Alió, C., and Acinas, S. G. (2012). Patterns and architecture of genomic

- islands in marine bacteria. *BMC Genomics* 13:347. doi: 10.1186/1471-2164-13-347
- Flynn, S., van Sinderen, D., Thornton, G. M., Holo, H., Nes, I. F., and Collins, J. K. (2002). Characterization of the genetic locus responsible for the production of ABP-118, a novel bacteriocin produced by the probiotic bacterium *Lactobacillus salivarius* subsp. *salivarius* UCC118. *Microbiology* 148(Pt 4), 973–984. doi: 10.1099/00221287-148-4-973
- Forde, B. M., Neville, B. A., O'Donnell, M. M., Riboulet-Bisson, E., Claesson, M. J., Coghlan, A., et al. (2011). Genome sequences and comparative genomics of two *Lactobacillus ruminis* strains from the bovine and human intestinal tracts. *Microb. Cell Fact.* 10(Suppl. 1):S13. doi: 10.1186/1475-2859-10-s1-s13
- Germond, J. E., Lapiere, L., Delley, M., and Mollet, B. (1995). A new mobile genetic element in *Lactobacillus delbrueckii* subsp. *bulgaricus*. *Mol. Gen. Genet.* 248, 407–416. doi: 10.1007/BF02191640
- Grissa, I., Vergnaud, G., and Pourcel, C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35, W52–W57. doi: 10.1093/nar/gkm360
- Harris, H. M. B., Bourin, M. J. B., Claesson, M. J., and Toole, P. W. (2017). Phylogenomics and comparative genomics of *Lactobacillus salivarius*, a mammalian gut commensal. *Microb. Genomics* 3:e000115. doi: 10.1099/mgen.0.000115
- Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., et al. (2016). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 44, D286–D293. doi: 10.1093/nar/gkv1248
- Husemann, P., and Stoye, J. (2010). r2cat: synteny plots and comparative assembly. *Bioinformatics* 26, 570–571. doi: 10.1093/bioinformatics/btp690
- Jimenez, E., Martin, R., Maldonado, A., Martin, V., Gomez de Segura, A., Fernandez, L., et al. (2010). Complete genome sequence of *Lactobacillus salivarius* CECT 5713, a probiotic strain isolated from human milk and infant feces. *J. Bacteriol.* 192, 5266–5267. doi: 10.1128/jb.00703-10
- Juhas, M., van der Meer, J. R., Gaillard, M., Harding, R. M., Hood, D. W., and Crook, D. W. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol. Rev.* 33, 376–393. doi: 10.1111/j.1574-6976.2008.00136.x
- Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44, D457–D462. doi: 10.1093/nar/gkv1070
- Kazou, M., Alexandraki, V., Pot, B., Tsakalidou, E., and Papadimitriou, K. (2017). Complete genome sequence of the dairy isolate *Lactobacillus acidipiscis* ACA-DC 1533. *Genome Announc.* 5:e01533-16. doi: 10.1128/genomeA.01533-16
- Khota, W., Pholsen, S., Higgs, D., and Cai, Y. (2016). Natural lactic acid bacteria population of tropical grasses and their fermentation factor analysis of silage prepared with cellulase and inoculant. *J. Dairy Sci.* 99, 9768–9781. doi: 10.3168/jds.2016-11180
- Kim, D. S., Choi, S. H., Kim, D. W., Kim, R. N., Nam, S. H., Kang, A., et al. (2011). Genome sequence of *Lactobacillus cypricasei* KCTC 13900. *J. Bacteriol.* 193, 5053–5054. doi: 10.1128/jb.05659-11
- Koyanagi, T., Nakagawa, A., Kiyohara, M., Matsui, H., Tsuji, A., Barla, F., et al. (2016). Tracing microbiota changes in yamahai-moto, the traditional Japanese sake starter. *Biosci. Biotechnol. Biochem.* 80, 399–406. doi: 10.1080/09168451.2015.1095067
- Lawson, P. A., Papademas, P., Wachter, C., Falsen, E., Robinson, R., and Collins, M. D. (2001). *Lactobacillus cypricasei* sp. nov., isolated from Halloumi cheese. *Int. J. Syst. Evol. Microbiol.* 51(Pt 1), 45–49. doi: 10.1099/00207713-51-1-45
- Li, N., Feng, F., Li, P., and Luo, L. X. (2014). Separation, identification, and characterization of food spoilage bacteria in soy sauce. *Mod. Food Sci. Technol.* 30, 120–125. doi: 10.13982/j.mfst.1673-9078.2014.10.021
- Li, P., Li, S., Cheng, L., and Luo, L. (2014). Analyzing the relation between the microbial diversity of DaQu and the turbidity spoilage of traditional Chinese vinegar. *Appl. Microbiol. Biotechnol.* 98, 6073–6084. doi: 10.1007/s00253-014-5697-4
- Li, Y., Raftis, E., Canchaya, C., Fitzgerald, G. F., van Sinderen, D., and O'Toole, P. W. (2006). Polyphasic analysis indicates that *Lactobacillus salivarius* subsp. *salivarius* and *Lactobacillus salivarius* subsp. *salicinii* do not merit separate subspecies status. *Int. J. Syst. Evol. Microbiol.* 56(Pt 10), 2397–2403. doi: 10.1099/ijs.0.64426-0
- Liu, M., Bayjanov, J. R., Renckens, B., Nauta, A., and Siezen, R. J. (2010). The proteolytic system of lactic acid bacteria revisited: a genomic comparison. *BMC Genomics* 11:36. doi: 10.1186/1471-2164-11-36
- Liu, M., Nauta, A., Francke, C., and Siezen, R. J. (2008). Comparative genomics of enzymes in flavor-forming pathways from amino acids in lactic acid bacteria. *Appl. Environ. Microbiol.* 74, 4590–4600. doi: 10.1128/aem.00150-08
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495. doi: 10.1093/nar/gkt1178
- Mao, P., Hu, Y., Liao, T., Wang, Z., Zhao, S., Liang, Y., et al. (2017). Microbial diversity during fermentation of sweet paste, a Chinese traditional seasoning, using PCR-denaturing gradient gel electrophoresis. *J. Microbiol. Biotechnol.* 27, 678–684. doi: 10.4014/jmb.1605.05011
- Martino, M. E., Bayjanov, J. R., Caffrey, B. E., Wels, M., Joncour, P., Hughes, S., et al. (2016). Nomadic lifestyle of *Lactobacillus plantarum* revealed by comparative genomics of 54 strains isolated from different habitats. *Environ. Microbiol.* 18, 4974–4989. doi: 10.1111/1462-2920.13455
- Melgar-Lalanne, G., Rivera-Espinoza, Y., Reyes Mendez, A. I., and Hernandez-Sanchez, H. (2013). In vitro evaluation of the probiotic potential of halotolerant lactobacilli isolated from a ripened tropical Mexican cheese. *Probiotics Antimicrob. Proteins* 5, 239–251. doi: 10.1007/s12602-013-9144-0
- Morales, F., Morales, J. I., Hernandez, C. H., and Hernandez-Sanchez, H. (2011). Isolation and partial characterization of halotolerant lactic acid bacteria from two Mexican cheeses. *Appl. Biochem. Biotechnol.* 164, 889–905. doi: 10.1007/s12010-011-9182-6
- Naser, S. M., Vancanneyt, M., Hoste, B., Snauwaert, C., and Swings, J. (2006). *Lactobacillus cypricasei* Lawson et al., 2001 is a later heterotypic synonym of *Lactobacillus acidipiscis* Tanasupawat et al., 2000. *Int. J. Syst. Evol. Microbiol.* 56(Pt 7), 1681–1683. doi: 10.1099/ijs.0.64229-0
- Neville, B. A., Forde, B. M., Claesson, M. J., Darby, T., Coghlan, A., Nally, K., et al. (2012). Characterization of pro-inflammatory flagellin proteins produced by *Lactobacillus ruminis* and related motile lactobacilli. *PLoS One* 7:e40592. doi: 10.1371/journal.pone.0040592
- Neville, B. A., and O'Toole, P. W. (2010). Probiotic properties of *Lactobacillus salivarius* and closely related *Lactobacillus* species. *Future Microbiol.* 5, 759–774. doi: 10.2217/fmb.10.35
- Nguyen, D. T. L., Van Hoorde, K., Cnockaert, M., De Brandt, E., De Bruyne, K., Le, B. T., et al. (2013). A culture-dependent and -independent approach for the identification of lactic acid bacteria associated with the production of nem chua, a Vietnamese fermented meat product. *Food Res. Int.* 50, 232–240. doi: 10.1016/j.foodres.2012.09.029
- Nierop Groot, M. N., and de Bont, J. A. M. (1998). Conversion of phenylalanine to benzaldehyde initiated by an aminotransferase in *Lactobacillus plantarum*. *Appl. Environ. Microbiol.* 64, 3009–3013.
- O'Shea, E. F., O'Connor, P. M., Raftis, E. J., O'Toole, P. W., Stanton, C., Cotter, P. D., et al. (2012). Subspecies diversity in bacteriocin production by intestinal *Lactobacillus salivarius* strains. *Gut Microbes* 3, 468–473. doi: 10.4161/gmic.21417
- Papadimitriou, K., Anastasiou, R., Mavrogonatou, E., Blom, J., Papandreou, N. C., Hamodrakas, S. J., et al. (2014). Comparative genomics of the dairy isolate *Streptococcus macedonicus* ACA-DC 198 against related members of the *Streptococcus bovis*/*Streptococcus equinus* complex. *BMC Genomics* 15:272. doi: 10.1186/1471-2164-15-272
- Perin, L. M., Savo Sardaro, M. L., Nero, L. A., Neviani, E., and Gatti, M. (2017). Bacterial ecology of artisanal Minas cheeses assessed by culture-dependent and -independent methods. *Food Microbiol.* 65(Suppl. C), 160–169. doi: 10.1016/j.fm.2017.02.005
- Pot, B., Felis, G. E., Bruyne, K. D., Tsakalidou, E., Papadimitriou, K., Leisner, J., et al. (2014). "The genus *Lactobacillus*," in *Lactic Acid Bacteria: Biodiversity and Taxonomy*, eds W. H. Holzappel and B. J. B. Wood (Hoboken, NJ: John Wiley & Sons), 249–353. doi: 10.1002/9781118655252.ch19
- Raftis, E. J., Forde, B. M., Claesson, M. J., and O'Toole, P. W. (2014). Unusual genome complexity in *Lactobacillus salivarius* JCM1046. *BMC Genomics* 15:771. doi: 10.1186/1471-2164-15-771
- Raftis, E. J., Salvetti, E., Torriani, S., Felis, G. E., and O'Toole, P. W. (2011). Genomic diversity of *Lactobacillus salivarius*. *Appl. Environ. Microbiol.* 77, 954–965. doi: 10.1128/aem.01687-10

- Randazzo, C. L., Todaro, A., Pino, A., Pitino, I., Corona, O., and Caggia, C. (2017). Microbiota and metabolome during controlled and spontaneous fermentation of Nocellara Etna table olives. *Food Microbiol.* 65(Suppl. C), 136–148. doi: 10.1016/j.fm.2017.01.022
- Reginensi, S. M., Olivera, J. A., Bermúdez, J., and González, M. J. (2016). “Lactobacillus in the dairy industry: from natural diversity to biopreservation resources,” in *Microbial Models: From Environmental to Industrial Sustainability*, ed. S. Castro-Sowinski (Singapore: Springer), 57–81. doi: 10.1007/978-981-10-2555-6_4
- Roberts, R. J., Vincze, T., Posfai, J., and Macelis, D. (2015). REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.* 43, D298–D299. doi: 10.1093/nar/gku1046
- Romeo, Y., Obis, D., Bouvier, J., Guillot, A., Fourcans, A., Bouvier, I., et al. (2003). Osmoregulation in *Lactococcus lactis*: BusR, a transcriptional repressor of the glycine betaine uptake system. *Mol. Microbiol.* 47, 1135–1147. doi: 10.1046/j.1365-2958.2003.03362.x
- Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics* 27, 1009–1010. doi: 10.1093/bioinformatics/btr039
- Sun, E., Ren, F., Liu, S., Ge, S., Zhang, M., Guo, H., et al. (2015). Complete genome sequence of *Lactobacillus salivarius* Ren, a probiotic strain with anti-tumor activity. *J. Biotechnol.* 20, 57–58. doi: 10.1016/j.jbiotec.2015.06.399
- 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
- Tanasupawat, S., Shida, O., Okada, S., and Komagata, K. (2000). *Lactobacillus acidipiscis* sp. nov. and *Weissella thailandensis* sp. nov., isolated from fermented fish in Thailand. *Int. J. Syst. Evol. Microbiol.* 50(Pt 4), 1479–1485. doi: 10.1099/00207713-50-4-1479
- Tanasupawat, S., Thongsanit, J., Okada, S., and Komagata, K. (2002). Lactic acid bacteria isolated from soy sauce mash in Thailand. *J. Gen. Appl. Microbiol.* 48, 201–209. doi: 10.2323/jgam.48.201
- Thamacharoensuk, T., Taweechotipatr, M., Kajikawa, A., Okada, S., and Tanasupawat, S. (2017). Induction of cellular immunity interleukin-12, antiproliferative effect, and related probiotic properties of lactic acid bacteria isolated in Thailand. *Ann. Microbiol.* 67, 511–518. doi: 10.1007/s13213-017-1280-4
- Tohno, M., Kobayashi, H., Nomura, M., Uegaki, R., and Cai, Y. (2012). Identification and characterization of lactic acid bacteria isolated from mixed pasture of timothy and orchardgrass, and its badly preserved silage. *Anim. Sci. J.* 83, 318–330. doi: 10.1111/j.1740-0929.2011.00955.x
- Tsuda, H., Kubota, K., Matsumoto, T., and Ishimi, Y. (2012). Isolation and identification of lactic acid bacteria in traditional fermented sushi, funazushi, from Japan. *Food Sci. Technol. Res.* 18, 77–82. doi: 10.3136/fstr.18.77
- van de Guchte, M., Penaud, S., Grimaldi, C., Barbe, V., Bryson, K., Nicolas, P., et al. (2006). The complete genome sequence of *Lactobacillus bulgaricus* reveals extensive and ongoing reductive evolution. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9274–9279. doi: 10.1073/pnas.0603024103
- van Heel, A. J., de Jong, A., Montalban-Lopez, M., Kok, J., and Kuipers, O. P. (2013). BAGEL3: automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. *Nucleic Acids Res.* 41, W448–W453. doi: 10.1093/nar/gkt391
- van Pijkeren, J. P., Canchaya, C., Ryan, K. A., Li, Y., Claesson, M. J., Sheil, B., et al. (2006). Comparative and functional analysis of sortase-dependent proteins in the predicted secretome of *Lactobacillus salivarius* UCC118. *Appl. Environ. Microbiol.* 72, 4143–4153. doi: 10.1128/aem.03023-05
- Varani, A. M., Siguier, P., Gourbeyre, E., Charneau, V., and Chandler, M. (2011). ISSaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. *Genome Biol.* 12:R30. doi: 10.1186/gb-2011-12-3-r30
- Visweswaran, G. R., Steen, A., Leenhouts, K., Szeliga, M., Ruban, B., Hesselting-Meinders, A., et al. (2013). AcmD, a homolog of the major autolysin AcmA of *Lactococcus lactis*, binds to the cell wall and contributes to cell separation and autolysis. *PLoS One* 8:e72167. doi: 10.1371/journal.pone.0072167
- Wanchai, V., Patumcharoenpol, P., Nookaew, I., and Ussery, D. (2017). dBBQs: dataBase of bacterial quality scores. *BMC Bioinformatics* 18(Suppl. 14):483. doi: 10.1186/s12859-017-1900-9
- Xie, Y., Wei, Y., Shen, Y., Li, X., Zhou, H., Tai, C., et al. (2018). TADB 2.0: an updated database of bacterial type II toxin-antitoxin loci. *Nucleic Acids Res.* 46, D749–D753. doi: 10.1093/nar/gkx1033
- Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. (2012). dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 40, W445–W451. doi: 10.1093/nar/gks479
- Yvon, M., and Rijnen, L. (2001). Cheese flavour formation by amino acid catabolism. *Int. Dairy J.* 11, 185–201. doi: 10.1016/S0958-6946(01)00049-8

Conflict of Interest Statement: 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.

Copyright © 2018 Kazou, Alexandraki, Blom, Pot, Tsakalidou and Papadimitriou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.