Brazilian Journal of Development

Comparative genomics study of *Lactobacillus plantarum* strains provides perspectives about proteins involved in the probiogenomics

Estudo de genômica comparativa de cepas de *Lactobacillus plantarum* fornece perspectivas sobre proteínas envolvidas com probiogenômica

Carlos Leonardo Araújo
Especialista em Microbiologia
Escola Superior da Amazônia – ESAMAZ
R. Arcipreste Manoel Teodoro, 820 - Batista Campos, Belém – PA
E-mail: leonardo.a.araujo@hotmail.com

Samuel Amaral
Especialista em Microbiologia
Escola Superior da Amazônia – ESAMAZ
R. Arcipreste Manoel Teodoro, 820 - Batista Campos, Belém – PA, Brasil
E-mail: samueltec2010@gmail.com

Tayná Novaes
Especialista em Microbiologia
Escola Superior da Amazônia – ESAMAZ
R. Arcipreste Manoel Teodoro, 820 - Batista Campos, Belém – PA, Brasil
E-mail: taynanovaes2@gmail.com

Géssica Faial
Especialista em Microbiologia
Instituição: Escola Superior da Amazônia – ESAMAZ
Endereço: R. Arcipreste Manoel Teodoro, 820 - Batista Campos, Belém – PA, Brasil
E-mail: gessicafaial@gmail.com

Larissa Dias
Doutora em Biotecnologia
Universidade do Estado do Pará – UEPA
Av. Plácido de Castro, 1399 – Aparecida, Santarém – PA, Brasil
E-mail: bio.larissadias@gmail.com

Ingrid Coutinho
Doutora em Agronomia
Escola Superior da Amazônia – ESAMAZ
R. Arcipreste Manoel Teodoro, 820 - Batista Campos, Belém – PA, Brasil
E-mail: ibernardodelima@gmail.com

Brazilian Journal of Development

ABSTRACT

Probiotic bacteria have received great attention in recent years due to the benefits conferred by this group of microorganisms on human health, such as the improvement of nutrient bioavailability, inhibiting the opportunistic pathogens growth and increasing host immune activity. In this context, the combination of *Lactobacillus plantarum* potential and the advent of NGS platforms contributed to the increase of genomes belonging to the species in the past decade. The present study aimed to use a series of bioinformatics tools to perform the comparative analysis of 41 completed genomes of the species. Genes involved in probiogenomics of the species were identified in the core genome, such as adhesion, antagonism to pathogens and host microbiota alterations. It was also detected a structural variation at genomic level, due to the identification of gene clusters present in certain strains and absent in others, a fact that may be related to the heterogeneity of habitats that *L. plantarum* can be isolated. Therefore, this study was able to provide new insights regarding the gene content of this important group of lactic acid bacteria, through the direct comparison of several strains, which opens the way for future studies aiming to understand the mechanisms by which *L. plantarum* is capable of causing its beneficial effects on human health.

Keywords: *Lactobacillus plantarum*, pan-genomics, phylogenomics, probiotics

RESUMO

As bactérias probióticas têm recebido grande atenção nos últimos anos devido aos benefícios conferidos por esse grupo de microrganismos à saúde humana, como a melhoria da biodisponibilidade de nutrientes, inibindo o crescimento de patógenos oportunistas e aumentando a atividade imune do hospedeiro. Nesse contexto, a combinação do potencial de *Lactobacillus plantarum* e o advento das plataformas NGS contribuíram para o aumento dos genomas pertencentes às espécies na última década. O presente estudo teve como objetivo utilizar uma série de ferramentas de bioinformática para realizar a análise comparativa de 41 genomas completos da espécie. Os genes envolvidos na probiogenômica da espécie foram identificados no genoma central, como adesão, antagonismo a patógenos e alterações da microbiota do hospedeiro. Também foi detectada variação estrutural no nível genômico, devido à identificação de agrupamentos de genes presentes em determinadas linhagens e ausentes em outras, fato que pode estar relacionado à heterogeneidade de habitats que *L. plantarum* pode ser isolada. Portanto, este estudo foi capaz de fornecer novos insights sobre o conteúdo gênico desse importante grupo de bactérias do ácido lático, através da comparação direta de várias cepas, o que abre caminho para estudos futuros, com o objetivo de entender os mecanismos pelos quais *L. plantarum* é capaz de causar seus efeitos benéficos na saúde humana.

Palavras-chave: *Lactobacillus plantarum*, pan-genoma, filogenômica, probióticos

1 INTRODUCTION

Bacteria have great importance for the food industry because they have a very diverse and rich enzymatic apparatus, which can be used in the meat, plants and milk fermentation [1-3]. In addition, they are important in the production of antimicrobial compounds that prevent the proliferation and dominance of pathogenic bacteria [4]. Also, they can be used as

Brazilian Journal of Development

probiotics, in which when ingested in small quantities they are able to confer benefits to the consumer health by providing vitamins and aiding in the digestion process [5].

In this context, the genus *Lactobacillus* is one of the most well-known nowadays, because it has remarkable probiotic properties. Despite the phylogenetic diversity, all species share the same capacity to produce lactic acid at the end of the fermentation process. The analysis of the variable region of the *rrs* (16S rRNA) gene is its main method of identification [6]. Within this genus, the species *Lactobacillus plantarum*, which aggregates non-pathogenic strains, is highlighted, being considered by the Food and Drug Administration (FDA) as GRAS organisms (Generally recognized as safe). They are heterofermentative bacteria, and can be naturally found in several niches, such as dairy products, meats, vegetables and the gastrointestinal tract humans and animals [7,8].

L. plantarum is a Gram-positive, non-sporulating and heterofermentative lactic acid bacterial species (BAL), being able to use several carbohydrate sources as substrate. It has rod shape and is generally about 0.9 to 1.2 μm x 3 to 8 μm long, occurring singly, in pairs or in short chains [9]. Several genomes have been sequenced in recent years, and present an average of 3.3 Mb, being one of the largest among the known BALs [8,10,11].

A recent phenotypic characterization of 185 strains obtained from different sources has demonstrated the existence of an enormous abundance and richness of this species within a certain niche [12]. The versatility and adaptive behavior of *L. plantarum* is a direct reflection of its genome, which exhibits an enormous amount of coding genes involved in the sugar entry and metabolization, thus allowing the use of a wide variety of carbohydrates available in the medium [13]. In addition, research has shown that these bacteria reduce the side effects caused by antibiotics and have a protective effect on the intestinal tract against the harmful effects induced by *Escherichia coli* [14,15].

One of the main features of this species is its ability to produce compounds with antimicrobial properties, including hydrogen peroxide and some organic acids and bacteriocins [16,17]. These metabolites avoid the pathogenic microorganisms' proliferation during the fermentation process. They also prevent the integration of harmful bacteria into the microbiota of the individual's intestines, therefore being important for the balance of ecological relationships in this environment [18].

One of the well-exploited attributes of probiotics is their ability to reduce fat absorption [19,20]. It is known that the composition of an individual's microbiota directly influences the deposition of this material by altering host metabolism [21,22]. *L. plantarum* strain K21, for

Brazilian Journal of Development

example, is able to inhibit the accumulation of lipids in preadipocytes, since an *in vivo* test demonstrated that this bacterium was able to delay the obesity of rats submitted to a diet high in fat [22].

Due to the advent of next-generation sequencing platforms (NGS) the number of genome designs in the three domains of life has increased exponentially in recent years, especially in the Bacteria domain. This large amount of data allowed the growth of bioinformatics, which provided the understanding of biological information previously unknown. Currently, several strains of the same phylogenetic clade can be compared at the genomic level, providing perspectives from the functional, metabolic and biochemical point of view of several bacteria of interest [23,24]. The pan-genomic approach, for example, consists of the comparative study of the repertoire of genes shared by all strains of a given population (core genome) and of those that are present only in some strains (accessory genome). The generated data are able to provide valuable information about the evolutionary and functional repertoire of a species [25].

Currently, there are more than 300 *L. plantarum* genomes deposited, thus enabling the evaluation of the complete genomic sequence present within the species. In this way, this work proposes to carry out a series of *in silico* analyzes in order to better understand the genomics of *L. plantarum* members, given the recognized probiotic potential of these bacteria.

2 MATERIAL AND METHODS

2.1 RETRIEVAL OF GENOME DATA

In this study, 41 complete genomes of *L. plantarum* available from the NCBI database were used. The strains were isolated from several sources in 12 countries. For this, two different file formats were obtained for each genome: FASTA and GBK. Subsequently, the nucleotide sequences were submitted to the RAST web platform [26] to standardize the files in order to avoid performance errors in the subsequent analyzes. The information regarding each strain used is contained in Table 1.

Table 1 - Genomic features of *L. plantarum* strains used in the study

Strain	Access NCBI	Genes	Isolation country	Isolation source
5--2	CP009236.1	3148	China	Fermented Soybean
10CH	CP023728.1	3192	United Kingdom	Cheese
b-2	CP027349.1	3120	China	Pickles
B21	CP010528.1	3189	Vietnam	Fermented sauce
BDGP2	CP023174.1	3559	United States	Feces
BLS41	CP018209.1	3401	China	Kimchi

Brazilian Journal of Development

CAUH2	CP015126.1	3188	China	Pickles
CGMCC 1.557	CP016270.1	3214	China	Uninformed
CLP0611	CP019722.1	3134	South Korea	Kimchi
DOMLa	CP004406.1	3103	Uninformed	Uninformed
DSR_M2	CP022294.1	3151	South Korea	Kimchi
GB-LP1	CP020564.1	2930	South Korea	Fermented Food
HFC8	CP012650.1	3433	India	Feces
JBE245	CP014780.1	3177	South Korea	Meju
JBE490	CP020861.1	3126	South Korea	Nuruk
JDM1	CP001617.1	3084	China	Uninformed
K25	CP020093.1	3335	Tibet	Kefir
K259	CP025590.1	3287	South Korea	Kimchi
KC28	CP026743.1	3232	South Korea	Kimchi
KC3	CP025586.1	3268	South Korea	Kimchi
KLDS1.0391	CP019348.1	2918	China	Jiaoke
LM1004	CP025988.1	3122	South Korea	Kimchi
LP2	CP020816.1	3185	China	Pickles
LP3	CP017066.1	3232	South Korea	Dietary Supplement
LQ80	CP028977.1	3371	Japan	Liquid Feed for Swine
LY-78	CP015308.1	3023	China	Sauerkraut
LZ227	CP015857.1	3393	China	Milk
LZ95	CP012122.1	3232	China	Feces
MF1298	CP013149.1	3489	Norway	Salami
P-8	CP005942.2	2952	China	Fermented Milk
RI-113	CP017406.1	3423	Switzerland	Salami
SRCM100434	CP021528.1	3097	South Korea	Uninformed
SRCM102022	CP021501.1	3261	South Korea	Uninformed
ST-III	CP002222.1	3194	China	Kimchi
TMW 1.1478	CP021932.1	3315	Germany	Honey
TS12	CP018324.1	3383	Malaysia	Tofu
WCFS1	AL935263.2	3174	United Kingdom	Human Saliva
X7021	CP025412.1	3336	China	Uninformed
Zhang-LL	CP011769.1	2860	China	Feces
ZJ316	CP004082.1	3209	China	Feces
ZS2058	CP012343.1	3077	China	Sauerkraut

2.2 PAN-GENOME ANALYSIS

The PanWeb platform was adopted for pan-genomic analysis [27]. PanWeb is an online interface based on the Pan Genome Analysis Pipeline - PGAP, a script in Perl language, developed by Zhao and collaborators (2012) [28]. For analysis of orthologous genes, we used alignment identity of 80%, coverage of 90% and e-value of 0.001. In addition, the pan-genome profile and orthologous distribution charts were constructed using the statistical package R [29], integrated into the program.

Brazilian Journal of Development

2.3 CONSTRUCTION OF THE PHYLOGENETIC TREE

The *rrs* gene sequences from each of the 41 strains were used for the construction of the phylogenetic tree. For the multiple alignment the BioEdit package was used, through the program ClustalW [30]. The evolutionary history was inferred using the Maximum Parsimony method by the MEGA7 program [31], in which the species *B. subtilis* ATCC 13952 was used as an external group.

2.4 PHYLOGENOMIC ANALYSIS

The comparison between the complete genomes was performed by the Gegenees program [32], which performs multiple alignments among the strains through the blastn algorithm, where each data between the genomes is represented by a score, that generates a heatmap. For the reliability of the results in this analysis the accurate parameter was used, which includes the fragments size of 200 basepairs, size of the reading space of 100 base and 30% of threshold.

2.5 CONSTRUCTION OF THE GENOMIC MAP

The comparison of the genomic architecture of the strains used was performed by BRIG software [33], with minimum and maximum threshold identities of 50% and 70%, respectively.

2.6 FUNCTIONAL ANNOTATION OF THE GENOME OF *L. PLANTARUM* JDM1

The main molecular functions, biological processes and cellular components of the orthologous genes had their respective roles investigated, using the Gene Ontology database by the GO FEAT program [34]. This platform has cross-references to the UniProt, InterPro, KEGG, Pfam, NCBI and SEED databases. *L. plantarum* strain JDM1 (CP001617.1) was used as reference genome, where the blastp algorithm was adopted with e-value 0.001 for the analysis performance.

3 RESULTS AND DISCUSSION

3.1 DISTRIBUTION OF ORTHOLOGOUS AND PAN-GENOME PROFILE

The orthologous analysis of the 41 *L. plantarum* genomes here analyzed showed that in this configuration the non-redundant gene repertoire of the species consists of about 8975 genes, of which 1580 constitute the core genome, 3900 are part of the accessory genome and

Brazilian Journal of Development

3495 are strain-specifics. The core genome is composed of genes that are shared by all the strains analyzed in the study. They are usually genes that encode products for basic aspects of the life, such as processes of replication, transcription, translation and maintenance of cellular homeostasis [35,36]. The accessory genome aggregates genes that are shared by some organisms, but not by all of them. In our analysis, the core genome of *L. plantarum* is composed of gene clusters present from two to 40 strains. Accessory genes possess varied functions, such as survival in specific environments, which are related to the life form of organisms [35].

In addition, the strains used in this work vary significantly in morphological and metabolic capacity terms, therefore, they can be used for different biotechnological purposes, being the most investigated their use in the food industry and in human health as probiotic [15,37]. For example, they are usually involved with the final fermentation stage because they are quite tolerant to acid environments [38,39].

The phenotypic variety of *L. plantarum* was observed by Siezen et al. (2010) in a study involving 185 isolates from different environment. They were characterized as to their fermentative behavior and growth. Isolates from the same group clustered together, except for the fecal samples, in which were distributed among the different groups, this allowed the researchers to infer that the origin of these strains is related to the food ingested by the individual. A deeper phenotypic analysis of 54 strains revealed that none of them was able to grow in milk at 4°C with high amount of nisin, due to the absence of genes that could confer survival in this medium. In contrast, the fermentation of certain sugars was restricted to only a few isolates, demonstrating the importance of advisory genes in natural selection [12]. In other study involving the genome comparison of the six strains it was possible to observe extremely conserved regions between them [40]. However, some variations were detected, especially in relation to the gene pools responsible for transposase and prophase coding. The prophages are derived from the interaction of bacteria with specific viruses, and their presence is more pronounced in pathogenic bacteria than those with free life [41].

The insertion of exogenous DNA may negatively interfere with the activity of an enzyme or it may be in charge of forming a new protein [42]. Transposases are enzymes responsible for the excision and transport of a DNA fragment known as a mobile transposable element to a given locus in the genome [43]. The genes encoding this enzyme are the most abundant in nature and are found in all living things. Like bacteriophages, transposases can override or activate a gene, which is why the gene involved in its synthesis is considered to be

Brazilian Journal of Development

selfish, one of its main characteristics being its propagation without concern for the integrity of the genome [44]. Another group of genes whose distribution was highly heterogeneous was involved with the synthesis of plantaricin, a bacteriocin composed of two peptides whose synergism promotes the inhibition of a huge variety of Gram-positive bacteria [45]. Bacteriocins differ from antibiotics because they have a higher specificity for their target and can act at nanoconcentration [46], and their biological role has been well investigated in *L. plantarum* [47].

Nwodo and colleagues also found the difference in the genes encoding proteins related to the synthesis of capsular or extracellular polysaccharides. Some parts of the gene cassette were common to the six isolates analyzed, while others were unique. This high variability reflects directly on the composition and formation process of the exopolysaccharides that exert important protective role [48]. The adhesion of some strains of *L. plantarum* in the intestine is mainly due to the action of this material. Its synthesis is important also in the biofilms formation that confer protection to the host against toxins and pathogenic bacteria [49].

Through the analysis to determine if a pan-genome is opened or closed, we can infer the number of genes and genomes that are necessary to identify the gene repertoire of the species, and this difference is related to the lifestyle of the bacteria studied [35]. When it is open, the pan genome needs a large amount of additional genes to characterize it; in contrast, in a closed pan-genome, bacteria have a low capacity to hold exogenous genes in their chromosomes, which is related to the isolated niches of which they are part, such as the endosymbionte *Buchnera aphidicola* [50]. It is possible to infer whether the pan-genome is opened or closed through the Heap Law ($n = k.N-\alpha$), where n is the expected number of genes from the pan-genome or core genome for a given number of N samples; k and α are values calculated according to the adjustment curve, which can be performed for the mean and median of each distribution. If the value of the element $\alpha \leq 1$ implies that the pan-genome is opened, but if $\alpha > 1$ the pan-genome is closed [27,51]. In this study, the value of α for 41 strains of L. *plantarum* is 0.6704 for the mean and 0.6697 for the median, implying that the pan-genome of the species is open (Figure 1).

𝔅razilian 𝒥ournal of 𝒟evelopment

Figure 1 - Pan-genome (blue) and core genome (red) profile of 41 *L. plantarum* strains. The plot shows the value of the Heap Law α coefficient when adjusting the curve for the mean (yellow line) and median (green line) of the different boxplots.
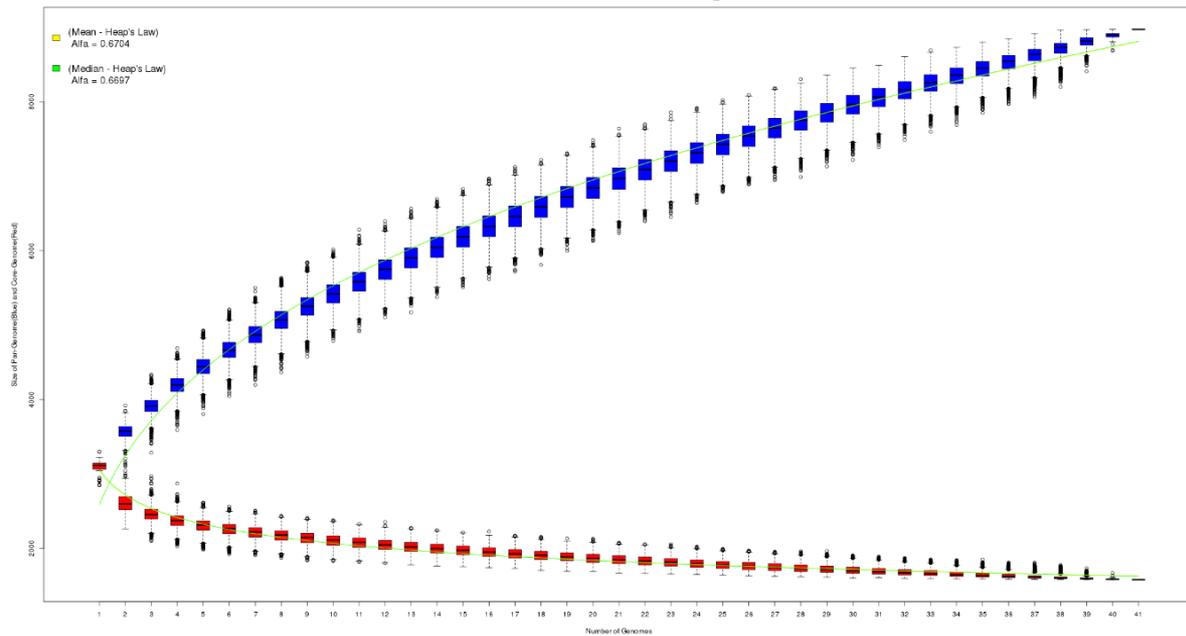

3.2 PHYLOGENETIC RELATIONSHIP AND FRAGMENTED ALIGNMENT OF THE STRAINS

Phylogenetic studies are important to describe the evolutionary history of species, being one of the best-known methods for the genetic differentiation of strains, especially through the analysis of the conserved and variable regions of the *rrs* gene, which encodes the 16S subunit of bacterial rRNA [6]. For example, a phylogenetic tree of *L. plantarum* isolates showed that other members of the species were clustered and distinctly separated from other strains of the genus *Lactobacillus*. The strain ZLP001 presented more than 99% similarity with other strains of *L. plantarum*, making it difficult to distinguish by the sequence similarity of the 16S rRNA gene [52]. The dendrogram based on the 16S rRNA sequence shows that the 41 strains analyzed in the present study also present a clonal-like behavior, presenting a high index of identity, based on the gene sequence.

Rivas et al. (2006) studied the population phylogeny of 16 isolates of L. *plantarum*, where the decomposition analysis indicated that recombination plays a role in the creation of genetic heterogeneity of the species [53], which reaffirms the tendency of members of the species have to carry significant changes in their genetic material. Interestingly, the tendency of high similarity between the 16S gene is not a constant, since when the analysis of the

Brazilian Journal of Development

complete genomes through the Gegenees was expanded, a different behavior was observed (Figure 2).

Figure 2: Analysis of the genomic composition of 41 strains of L. plantarum. To the left is shown the dendrogram, based on the sequences of the rrs genes, and to the right is shown the heatmap of the comparison of each pair of genomes based on the similarity of fragments, where the position of the strains in the vertical corresponds to the numbers in the horizontal. The color scale of the graphs reflects the percentage of identity, ranging from the smallest (red) to the largest (green).
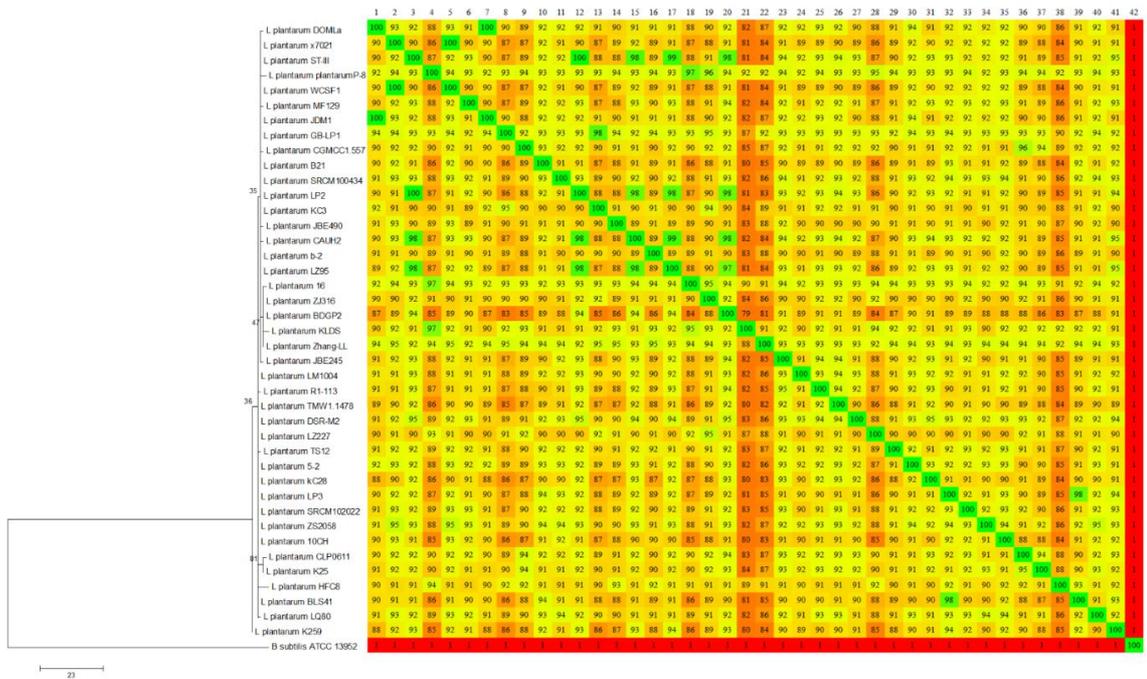

In relation to the genomic difference among the strains, comparative analyzes involving members of the subgroup *Lactobacillus casei* (*L. casei, L. paracasei* and *L. rhamnosus*) indicated a wide distribution of prophages in the analyzed bacteria. The impact of these phage-host interactions on the probiotic, phenotypic and stability properties of *Lactobacillus* strains indicate that the presence of bacteriophage can also contribute to the high frequency of recombination events with their hosts, leading to the genotypic diversity of both [54].

3.3 GENOMIC ARCHITECTURE OF *L. PLANTARUM*

In Figure 3 it is possible to observe the presence of at least 12 regions that are present in some strains and absent in others. These findings corroborate our results, since in the analysis of homologues about 82.3% of the repertoire of analyzed genes do not present homologues in the 41 studied strains. Horizontal transference of genes is a phenomenon that is well studied in bacteria, and the fact that *L. plantarum* is considered a free-living

Brazilian Journal of Development

microorganism facilitates the exchange of genetic material frequently, due to the diversity of environments and sources of isolation in which strains used in our analyzes are based [5].

Figure 3 - Genomic map of *L. plantarum* strains used in the study. The colored rings represent the 41 complete genomes used. In the center, the black lines represent the GC content, the purple and green lines represent the GC skew.
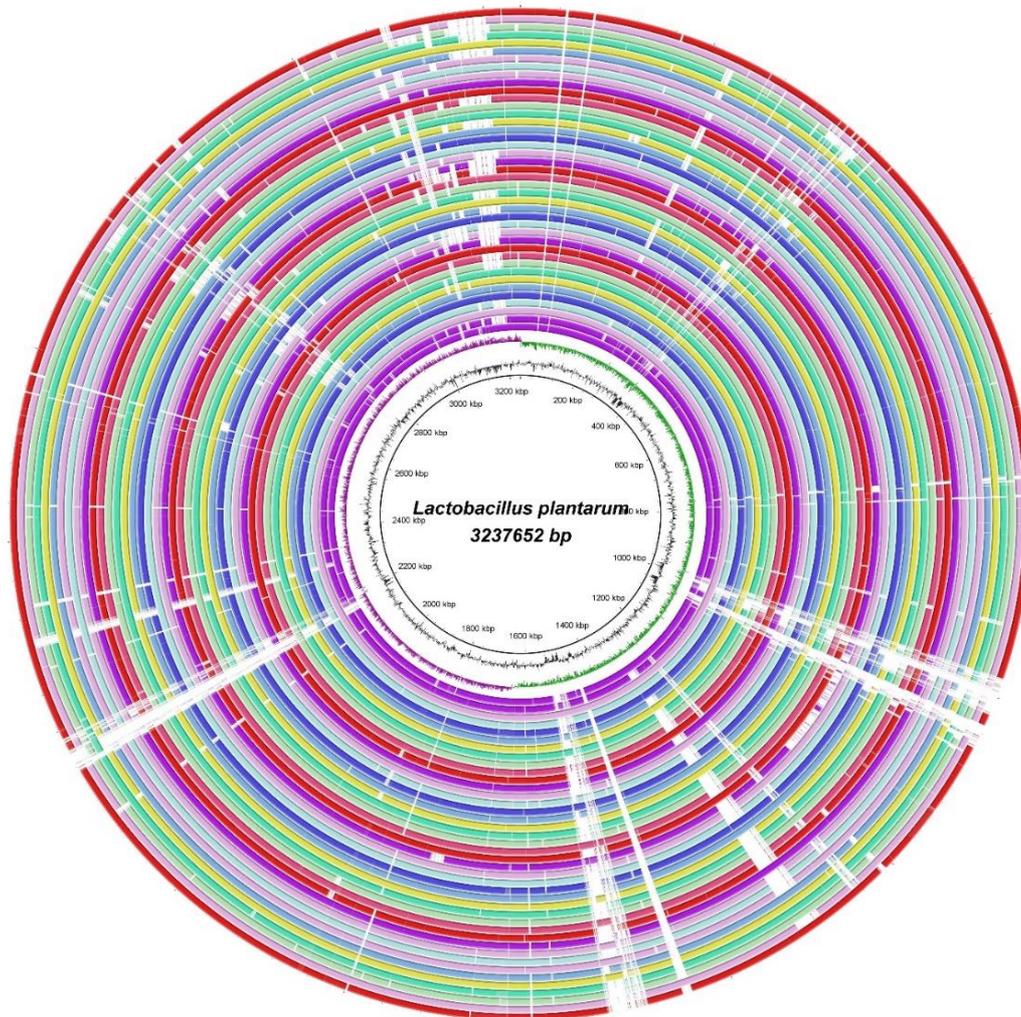
3.4 PREDICTION OF FUNCTIONAL CATEGORIES

Genes involved with carbohydrate metabolism were the majorities in this analysis. It is known that *L. plantarum* has an excellent adaptive capacity. The first genome of the species to be sequenced was from the WCFS1 strain and revealed the presence of one of the largest chromosomes detected at the level of genus and three plasmids. A great amount of genes coding for proteins linked to the capture, transport and utilization of sugars was found. Thus, the high survival capacity of this microorganism was justified in the most varied environments [56, 57].

Brazilian Journal of Development

3.5 PROBIOGENOMICS OF *L. PLANTARUM*

The functional annotation also provided clues about the genes involved with the probiotic potential of the species, some of which are already well known in the literature. Despite this, the molecular mechanisms of how these bacteria positively impact host health have not yet been fully elucidated. For this reason, these microorganisms represent a growing area of research that involves concepts of microbiology, genomics, molecular biology and genetics, known as probiogenomics [58]. The integration of probiogenomic and functional information with data on the expression of the host gene in the human intestine broadens the understanding of the roles of intestinal microbiota interactions, between microorganisms and between hosts and microorganisms. This integration of approaches provides the simultaneous analysis of a large volume of genes and proteins [59].

Among the genes related to probiotic potential, some with bacteriocin activity in the *L. plantarum* were predicted, three of which are also present in the coregenome (UniProt IDs A0A0G9GSP3, A0A199QDK6 and A0A199QE21). Bacteriocins are peptides produced by bacteria that act against other microorganisms whose producer has mechanisms of immunity, being able to inhibit the growth of undesirable organisms, and can aid in probiotic functionality through three mechanisms: i) helping the survival of production line in the intestinal environment; ii) inhibition of pathogen growth and iii) acting as quorum signaling peptides in the intestinal environment [60].

The ability of adhesion to epithelial tissue of the host's gastrointestinal tract and colonization is also another desired characteristic in probiotic bacteria. The adhesion of lactobacilli to intestinal cells has shown that the factors involved in the interaction vary between different species and strains, as well as the interaction between bacterial adhesins and the extracellular matrix or mucus proteins. The protein enolase-1 predicted in this study (UniProt ID Q88YH3) has been reported as the factor that adheres to fibronectin, a glycoprotein that binds to cells, collagen and glycosaminoglycans, helping to mediate and maintain cell migration and adhesion intestine [61].

The mechanism that BALs possess to hydrolyze bile salts has often been included as one of the criteria for the selection of probiotic strains, and several Bile Salt Hydrolase (BSH) hydrolases have been identified and characterized 62, including one present in all strains used in the comparative analyzes of this study (UniProt ID Q88SP0).

These enzymes are capable of altering the composition of the intestinal microbiota and metabolism, detoxifying the bile, in addition to lowering cholesterol levels. Bile acids play a

Brazilian Journal of Development

key role in the formation of intestinal microbiota and host physiology, regulating metabolic signaling, weight gain, as well as serum cholesterol levels and liver triglycerides [63]. In this way, species that contain BSH can act in a positive way in the reduction of circulating cholesterol levels, since with the desconjugation of the bile salts in the intestine, the presence of these salts decreases, inducing the production of bile from the cholesterol. Moreover, from an evolutionary point of view, these bacteria are mainly associated with nests adapted to vertebrates, demonstrating selective pressure on BALs to evolve and adapt to specific environments [62,64].

4 CONCLUSION

The technological advances provided by large-scale genome sequencing have enabled the scientific community to gain a greater understanding of genetic, metabolic, biochemical and adaptive mechanisms in all three domains of life. The DNA sequences of bacteria of medical, industrial and environmental interest are available in databases, which facilitated the execution of research that compare the genetic content of these microorganisms.

Through this study, it was possible to compare, through a robust *in silico* strategies methodology, the complete genomes of 41 strains of *L. plantarum*, an important member of the BALs group. The pan-genome of the bacterium consists of more than 8000 genes, where some of which have orthologs in all isolates are notorious for their probiotic activities, such as adhesion, antagonism to pathogens and alteration of the host microbiota. It was also detected a structural variation at genomic level, due to the identification of gene clusters present in certain strains and absent in others, a fact that may be related to the heterogeneity of habitats from which *L. plantarum* can be isolated.

Therefore, this study was able to provide new insights regarding the gene content of this important microorganism through the direct comparison of several strains, which opens the way for future studies aiming to understand the mechanisms by which *L. plantarum* is capable of causing its beneficial effects on human health.

Brazilian Journal of Development

REFERENCES

1.   Guo W, Jia W, Li Y, Chen S (2010). Performances of Lactobacillus brevis for Producing Lactic Acid from Hydrolysate of Lignocellulosics. Appl Biochem Biotechnol 161:124 – 136. https://doi.org/10.1007/s12010-009-8857-8

2.   E. Mani-López, E. Palou, A. López-Malo (2014). Probiotic viability and storage stability of yogurts and fermented milks prepared with several mixtures of lactic acid bacteria. J Dairy Sci 97(5):2578 – 257890. http://dx.doi.org/10.3168/jds.2013-7551

3.   Reddy LV, Park J, Wee Y (2015). Homofermentative Production of Optically Pure L-lactic Acid from Sucrose and Mixed Sugars by Batch Fermentation of Enterococcus faecalis RKY1. Biotechnol. Bioprocess Eng 20(6):1099 – 1105. http://dx.doi.org/10.1007/s12257-015-0379-3

4.   Perez RH, Zendo T, Sonomoto K (2014). Novel bacteriocins from lactic acid bacteria (LAB): various structures and applications. Microb Cell Fact. 13(Suppl 1):S3. http://dx.doi.org/10.1186/1475-2859-13-S1-S3

5.   Campana R, Hemert S Van, Baffone W (2017). Strain‑specific probiotic properties of lactic acid bacteria and their interference with human intestinal pathogens invasion. Gut Pathog 6:1 – 12. http://dx.doi.org/ 10.1186/s13099-017-0162-4

6.   Mahato NK, Gupta V, Singh P, Kumari R, Verma H, Tripathi C, et al (2017). Microbial taxonomy in the era of OMICS: application of DNA sequences, computational tools and techniques.   Antonie   Van   Leeuwenhoek   110(10):1357   –   1371. http://dx.doi.org/10.1007/s10482-017-0928-1

7.   Valerio F, Bellis P De, Lonigro SL, Visconti A, Lavermicocca P (2008). Use of Lactobacillus plantarum fermentation products in bread-making to prevent Bacillus subtilis ropy   spoilage.   Int   J   Food   Microbiol   2008;122:328   –   332. http://dx.doi.org/10.1016/j.ijfoodmicro.2008.01.005

8.   Jia F, Zhang L, Pang X, Gu X, Abdelazez A, Liang Y (2017). Complete genome sequence of bacteriocin-producing Lactobacillus plantarum KLDS1.0391, a probiotic strain with gastrointestinal tract resistance and adhesion to the intestinal epithelial cells. Genomics 109(5–6):432 – 437. https://dx.doi.org/10.1016/j.ygeno.2017.06.008

9.   Landete JM, Rodríguez H, Curiel JA, Rivas B De, De FL (2010). Degradation of Phenolic Compounds Found in Olive Products by Lactobacillus plantarum Strains. In: Olives and Olive Oil in Health and Disease Prevention. Elsevier Inc.; 2010. p. 387 – 396. http://dx.doi.org/10.1016/B978-0-12-374420-3.00043-7

10.   Li P, Zhou Q, Gu Q (2016). Complete genome sequence of Lactobacillus plantarum LZ227, a potential probiotic strain producing B-group vitamins. J Biotechnol 234:66 – 70. http://dx.doi.org/10.1016/j.jbiotec.2016.07.020

11.   Altermann E, Russell WM, Azcarate-Peril MA, Barrangou R, Buck BL, Mcauliffe O, et al (2005). Complete genome sequence of the probiotic lactic acid bacterium Lactobacillus acidophilus NCFM. Proc Natl Acad Sci USA 15;102(11):3906 – 3912. http://dx.doi.org/10.1073/pnas.0409188102

12.   Siezen RJ, Tzeneva VA, Castioni A, Wels M, Phan HTK, Rademaker JLW, et al (2010). Phenotypic and genomic diversity of Lactobacillus plantarum strains isolated from various environmental niches. Environ Microbiol 12(3):758 – 773. http://dx.doi.org/10.1111/j.1462-2920.2009.02119.x

13.   Kleerebezem M, Boekhorst J, Kranenburg R Van, Molenaar D, Kuipers OP, Leer R, et al (2003). Complete genome sequence of Lactobacillus plantarum WCFS1. Proc Natl Acad Sci USA 18;100(4):1990 – 1995. http://dx.doi.org/10.1073/pnas.0337704100

14.   Dhanani AS, Bagchi T (2013). Lactobacillus plantarum CS24.2 prevents Escherichia coli adhesion to HT-29 cells and also down-regulates enteropathogen-induced tumor necrosis factor-α and interleukin-8 expression. 2013. Microbiol Immunol 57(4):309 – http://dx.doi.org /315. 10.1111/1348-0421.12038

15.   Behera SS, Ray RC, Zdolec N (2018). Lactobacillus plantarum with Functional Properties: An Approach to Increase Safety and Shelf-Life of Fermented Foods. Biomed Res Int 2018: 1 – 19. https://dx.doi.org/10.1155/2018/9361614

16.   Neal-McKinney JM, Lu X, Duong T, Larson CL, Call DR, Shah DH, et al (2012). Production of Organic Acids by Probiotic Lactobacilli Can Be Used to Reduce Pathogen Load in Poultry. PLoS One 7(9):1 – 11. https://dx.doi.org/10.1371/journal.pone.0043928

17.   Cortés-Zavaleta O, López-Malo A, Hernández-Mendoza A, García HS (2014). Antifungal activity of lactobacilli and its relationship with 3-phenyllactic acid production. Int J Food Microbiol 173:30 – 35. http://dx.doi.org/10.1016/j.ijfoodmicro.2013.12.016

18.   Sabo S da S, Vitolo M, González JMD, Oliveira RP de S (2014). Overview of Lactobacillus plantarum as a promising bacteriocins producer among lactic acid bacteria. Food Res Int 64:527 – 536. http://dx.doi.org/10.1016/j.foodres.2014.07.041

19.   Park S, Seong K, Lim S (2016). Anti-obesity Effect of Yogurt Fermented by Lactobacillus plantarum Q180 in Diet-induced Obese Rats. Korean J Food Sci Anim Resour 36(1): 77 – 83. http://dx.doi.org/10.5851/kosfa.2016.36.1.77

20.   Kim S, Huang E, Park S, Holzapfel W (2018). Physiological Characteristics and Anti-obesity Effect of Lactobacillus plantarum K10. Korean J Food Sci Anim Resour 38(3):554 – 569. http://dx.doi.org/10.5851/kosfa.2018.38.3.554

21.   Turnbaugh PJ, Bäckhed F, Fulton L, Gordon JI (2008). Diet-Induced Obesity Is Linked to Marked but Reversible Alterations in the Mouse Distal Gut Microbiome. Cell Press 3(4):213 – 223. https://dx.doi.org/10.1016/j.chom.2008.02.015

22.   Wu C-C, Weng W-L, Lai W-L, Tsai H-P, Liu W-H, Lee M-H, et al (2015). Effect of Lactobacillus plantarum Strain K21 on High-Fat-Diet-Fed Obese Mice. Evidence-Based Complement Altern Med. 2015:1 – 9. http://dx.doi.org/10.1155/2015/391767

23.   Van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C (2018). The Third Revolution in Sequencing Technology. Trends Genet 34(9):666 - 681. http://dx.doi.org/10.1016/j.tig.2018.05.008

24.   Vincent AT, Derome N, Boyle B, Culley AI, Charette SJ (2017). Next-generation sequencing (NGS) in the microbiological world: How to make the most of your money. J Microbiol Methods 138:60 – 71. http://dx.doi.org/10.1016/j.mimet.2016.02.016

25.   Golicz AA, Batley J, Edwards D (2016). Towards plant pangenomics. Plant Biotechnol J 14(4):1099 – 1105. http://dx.doi.org/10.1111/pbi.12499

26.   Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al (2014). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42(D1):206 – 214. https://dx.doi.org/ 10.1093/nar/gkt1226

27.   Pantoja Y, Pinheiro K, Veras A, Araújo F, Lopes de Sousa A, Guimarães LC, et al (2017). PanWeb: A web interface for pan-genomic analysis. PLoS One 12(5):e0178154. http://dx.doi.org/10.1371/journal.pone.0178154

28.   Zhao Y, Wu J, Yang J, Sun S, Xiao J, Yu J (2012). PGAP: Pan-genomes analysis pipeline. Bioinformatics 28(3):416 – 418. http://dx.doi.org/10.1093/bioinformatics/btr655

29.   Snipen L, Liland KH (2015). micropan: an R-package for microbial pan-genomics. BMC Bioinformatics. 16(79):1 – 8. https://dx.doi.org/10.1186/s12859-015-0517-0

30.   Thompson JD, Higgins DG, Gibson TJ (1994). ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res 22(22):4673 – 4680

31.   Kumar S, Stecher G, Tamura K (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33(7):1870 – 1874. https://dx.doi.org/10.1093/molbev/msw054

32.   Ågren J, Sundström A, Hafström T, Segerman B (2012). Gegenees: Fragmented Alignment of Multiple Genomes for Determining Phylogenomic Distances and Genetic Signatures Unique for Specified Target Groups. PLoS One 7(6):e39107. https://dx.doi.org/10.1371/journal.pone.0039107

33.   Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics 12(1):402 1 – 10. https://dx.doi.org/10.1186/1471-2164-12-402

34.   Araujo FA, Barh D, Silva A, Guimarães L, Ramos RTJ (2018). GO FEAT: a rapid web-based functional annotation tool for genomic and transcriptomic data. Sci Rep 8(1):1794. https://dx.doi.org/10.1038/s41598-018-20211-9

35.   Medini D, Donati C, Tettelin H, Masignani V, Rappuoli R (2005). The microbial pan-genome. Curr Opin Genet Dev 15(6):589 – 594. https://dx.doi.org/10.1016/j.gde.2005.09.006

36.   Guimarães LC, de Jesus LB, Viana MVC, Silva A, Ramos RTJ, Soares S de C, et al (2015). Inside the Pan-genome - Methods and Software Overview. Curr Genomics 2015;16(4):245 – 252. https://dx.doi.org/10.2174%2F1389202916666150423002311

37.   Guidone A, Zotta T, Ross RP, Stanton C, Rea MC, Parente E, et al (2014). Functional properties of Lactobacillus plantarum strains: A multivariate screening study. Food Sci Technol 2014;56(1):69 – 76.http://dx.doi.org/10.1016/j.lwt.2013.10.036

38.   Plengvidhya V, Breidt F, Lu Z, Fleming HP (2007). DNA Fingerprinting of Lactic Acid Bacteria in Sauerkraut Fermentations Appl Environ Microbiol 73(23):7697 – 702. https://dx.doi.org/10.1128/AEM.01342-07

39.   Luxananil P, Promchai R, Wanasen S, Kamdee S, Thepkasikul P, Plengvidhya V, et al (2009). Monitoring Lactobacillus plantarum BCC 9546 starter culture during fermentation of Nham, a traditional Thai pork sausage. Int J Food Microbiol 2009;129(3):312 – 315. http://dx.doi.org/10.1016/j.ijfoodmicro.2008.12.011

40.   Siezen RJ, Et J, Vlieg VH (2011). Genomic diversity and versatility of Lactobacillus plantarum, a natural metabolic engineer. Microb Cell Fact 2011;10(Suppl 1):S3. https://dx.doi.org/10.1186%2F1475-2859-10-S1-S3

41.   Canchaya C, Fournous G, Brüssow H (2004). The impact of prophages on bacterial chromosomes. Mol Microbiol 53(1):9 – 18. https://dx.doi.org/10.1111/j.1365-2958.2004.04113.x

42.   Lee CY, Iandolo JJ (1986). Lysogenic conversion of staphylococcal lipase is caused by insertion of the bacteriophage L54a genome into the lipase structural gene. J Bacteriol 166(2):385 – 391.

43.   Dubin MJ, Scheid OM, Becker C (2018). Transposons: a blessing curse. Curr Opin Plant Biol 42:23 – 29. https://dx.doi.org/10.1016/j.pbi.2018.01.003

44.   Aziz RK, Breitbart M, Edwards RA (2010). Transposases are the most abundant, most ubiquitous genes in nature. Nucleic Acids Res 2010;38(13):4207 – 4217. https://dx.doi.org/10.1093/nar/gkq140

45.   Holo H, Jeknic Z, Daeschel M, Stevanovic S, Nes IF (2001). Plantaricin W from Lactobacillus plantarum belongs to a new family of two-peptide lantibiotics. Microbiology 147:643–51. https://dx.doi.org/10.1099/00221287-147-3-643

46.   Cotter PD, Ross RP, Hill C (2012). Bacteriocins - a viable alternative to antibiotics? Nat Rev Microbiol 11(2):95–105. http://dx.doi.org/10.1038/nrmicro2937

47.   Todorov SD (2009). Bacteriocins from Lactobacillus plantarum – Production, Genetic Organization and mode of action. Brazilian J Microbiol. 40:209 – 221. https://dx.doi.org/10.1590%2FS1517-83822009000200001

48.   Nwodo UU, Green E, Okoh AI (2012). Bacterial Exopolysaccharides: Functionality and Prospects. 13(11):14002 – 14015. https://dx.doi.org/10.3390/ijms131114002

49.   Castro-Bravo N, Wells JM, Margolles A, Ruas-Madiedo P (2018). Interactions of Surface Exopolysaccharides From Bifidobacterium and Lactobacillus Within the Intestinal Environment. Front Microbiol 9:1 – 15. https://dx.doi.org/10.3389/fmicb.2018.02426

50.   Vernikos G, Medini D, Riley DR, Tettelin H (2015). Ten years of pan-genome analyses. Curr Opin Microbiol 23:148 – 154. https://dx.doi.org/10.1016/j.mib.2014.11.016

51.   Tettelin H, Riley D, Cattuto C, Medini D (2008). Comparative genomics: the bacterial pan-genome.   Curr   Opin   Microbiol   11(5):472   –   477. https://dx.doi.org/10.1016/j.mib.2008.09.006

52.   Zhang W, Ji H, Zhang D, Liu H, Wang S, Wang J, et al (2018). Complete Genome Sequencing of Lactobacillus plantarum ZLP001, a Potential Probiotic That Enhances Intestinal Epithelial Barrier Function and Defense Against Pathogens in Pigs. Front Physiol 9:1 – 7. https://dx.doi.org/10.3389/fphys.2018.01689

53.   De las Rivas B, Marcobal Á, Muñoz R (2006). Development of a multilocus sequence typing method for analysis of Lactobacillus plantarum strains. Microbiology 152:85 – 93. https://dx.doi.org/10.1099/mic.0.28482-0

54.   Mercanti DJ, Rousseau GM, Capra ML, Quiberoni A, Tremblay DM, Labrie SJ (2016). Genomic Diversity of Phages Infecting Probiotic Strains of Lactobacillus paracasei. Appl Environ Microbiol 82(1):95 – 105. https://dx.doi.org/10.1128/AEM.02723-15

55.   Soucy SM, Huang J, Gogarten JP (2015). Horizontal gene transfer: building the web of life. Nat Rev Genet 16(8):472 – 482. http://dx.doi.org/10.1038/nrg3962

56.   Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al (2000). Gene Ontology: Tool for the unification of biology. Nat Genet 25(1):25 – 29. https://dx.doi.org/10.1038/75556

57.   Watanabe M, Van der Veen S, Abee T (2012). Impact of Respiration on Resistance of Lactobacillus plantarum WCFS1 to Acid Stress. Appl Environ Microbiol 78(11):4062 – 4064. https://dx.doi.org/10.1128/AEM.00287-12

58.   Ventura M, Turroni F, Van Sinderen D (2012). Probiogenomics as a tool to obtain genetic insights into adaptation of probiotic bacteria to the human gut. Bioeng Bugs 3(2):73 – 79. https://dx.doi.org/10.4161/bbug.18540

Brazilian Journal of Development

59.   Ventura M, Flaherty SO, Claesson MJ, Turroni F, Klaenhammer D, Van Sinderen D, O'Toole PW (2009). Genome-scale analyses of health-promoting bacteria: probiogenomics. Nat Rev Microbiol 7(1): 61 – 71. https://dx.doi.org/10.1038/nrmicro2047

60.   Hegarty JW, Guinane CM, Ross RP, Hill C, Cotter PD (2016). Bacteriocin production: a relatively unharnessed probiotic trait?. F1000Res 5: 2587. https://dx.doi.org/10.12688%2Ff1000research.9615.1

61.   Castaldo C, Vastano V, Siciliano RA, Candela M, Vici M, Muscariello L et al (2009). Surface displaced alfa-enolase of Lactobacillus plantarum is a fibronectin binding protein. Microb Cell Fact 8:14. https://dx.doi.org/10.1186/1475-2859-8-14

62.   Begley M, Hill C, Gahan CGM (2006). Bile Salt Hydrolase Activity in Probiotics. Appl Environ Microbiol 72(3):1729 – 1738. https://dx.doi.org/10.1128/AEM.72.3.1729-1738.2006

63.   Halder D, Mandal M, Chatterjee S, Pal N, Mandal S (2017). Indigenous Probiotic Lactobacillus Isolates Presenting Antibiotic like Activity against Human Pathogenic Bacteria. Biomedicines 5(2):31. https://dx.doi.org/10.3390/biomedicines5020031

64.   O'Flaherty S, Crawley AB, Theriot CM, Barrangou R (2018). The Lactobacillus Bile Salt Hydrolase Repertoire Reveals Niche-Specific Adaptation. mSphere. 3(3):1 – 13. https://dx.doi.org/10.1128/mSphere.00140-18