

TITLE: COMPARATIVE GENOMICS STUDY OF *LACTOBACILLUS PLANTARUM* STRAINS PROVIDES PERSPECTIVES ABOUT PROTEINS RELATED TO PROBIOGENOMICS

AUTHORS: ARAÚJO, C.L.A.^{1, 2}; AMARAL, S.C.^{1, 2}; NOVAES, T.I.C.¹; FAIAL, G.B.¹; DIAS, L.M.²; LIMA, I.B.¹.

INSTITUTION: ¹ ESPECIALIZAÇÃO *LATO SENSU* EM MICROBIOLOGIA, ESCOLA SUPERIOR DA AMAZÔNIA – ESAMAZ (RUA ARCIPRESTE MANOEL TEODORO, 820 – BATISTA CAMPOS, BELÉM - PA, 66015-040); ² UNIVERSIDADE FEDERAL DO PARÁ – UFPA (RUA AUGUSTO CORRÊA, 1 – GUAMÁ, BELÉM – PA, 66075-110).

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 improve of nutrient bioavailability, inhibiting the growth of opportunistic pathogens and increasing host immune activity. In this context, the combination of *Lactobacillus plantarum* potential and the advent of NGS platforms provided 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 complete genomes of the species. At first, the genome sequences were standardized by the RAST database, followed by the analysis of orthologous and unique genes through the PanWeb platform, adopting 80% identity, 90% coverage, and 0.001 *e-value*. The pan-genome profile plot was constructed through by the R statistical package. A phylogenetic tree based on the 16 rRNA gene sequences was built using the Maximum Parsimony method by the MEGA7 program, in which the *Bacillus subtilis* ATCC 13952 was used as an external group. The comparison between the complete genomes was performed by the Gegenees, and the BRIG tool provided the genomic architecture comparison from the strains. The pan-genome of *L. plantarum* consists in 8975 genes, where 1580 are part of the core genome. The phylogenetic analysis showed little genetic variety of the isolates, but at the genomic level, the strains presented significant differences. These findings corroborate the previous results, since in the orthologous analysis, about 82.3% of the gene repertoire analyzed do not present homologues in the 41 strains. Besides that, 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.