



Ministério da educação

Universidade Federal do Rio Grande

Programa de Pós-Graduação em Ciências da Saúde



**PROSPECÇÃO *IN SILICO* DE CEPAS DE *LACTOBACILLUS ACIDOPHILUS* COM
POTENCIAL ATIVIDADE PROBIÓTICA**

Rafaella Sinnott Dias

Rio Grande, 2023



Ministério da educação

Universidade Federal do Rio Grande

Programa de Pós-Graduação em Ciências da Saúde



**PROSPECÇÃO *IN SILICO* DE CEPAS DE *LACTOBACILLUS ACIDOPHILUS* COM
POTENCIAL ATIVIDADE PROBIÓTICA**

Rafaella Sinnott Dias

Dissertação apresentada ao Programa de Pós-graduação em Ciências da Saúde da Universidade Federal do Rio Grande, como requisito parcial à obtenção do título de Mestre em Ciências da Saúde

Orientadora: Prof. Dr^a Luciana Farias da Costa de Avila

Co-orientador: Prof. Dr. Frederico Schmitt Kremer

Rafaella Sinnott Dias

Dissertação apresentada ao
Programa de Pós-graduação em
Ciências da Saúde da Universidade
Federal do Rio Grande, como
requisito parcial à obtenção do
título de Mestre em Ciências da
Saúde

**PROSPECÇÃO *IN SILICO* DE CEPAS DE *LACTOBACILLUS ACIDOPHILUS* COM
POTENCIAL ATIVIDADE PROBIÓTICA**

Banca Examinadora:

Prof. Dra. Daniela Fernandes Ramos (Membro interno PPGCS)

Prof. Dr. Luciano Pinto (Membro externo UFPEL)

Prof. Dr. Carlos James Scaini (Membro suplente PPGCS)

Prof. Dra. Luciana Farias da Costa de Avila (Orientadora)

Prof. Dr. Frederico Schmitt Kremer (Co-orientador)

AGRADECIMENTOS

Agradeço a minha mãe, Tânia, uma das mulheres mais fortes que conheci. Ela que apesar de estar passando por um momento difícil, sempre foi a minha base, mesmo em momentos que eu deveria ser a dela, me motivando a acreditar que eu conseguiria concluir essa etapa, dizendo-me a seguinte frase: “minha filha, tenho certeza que és capaz”. Agradeço também a minha irmã Antonella, por muitas vezes cuidar da minha mãe enquanto ela estava em quimioterapia, proporcionando que eu ficasse estudando e trabalhando no meu projeto.

Agradeço aos meus amigos do PPGCS que foram meu verdadeiro porto seguro nesta trajetória, obrigada por tornarem tão leve essa jornada. Agradeço ao meu “bichinho de estimação” Maya, ela que foi minha companheira de longas horas de estudo, que me ouvia ensaiar apresentações e que sempre foi minha companheira fiel.

Agradeço ao Programa de Pós-Graduação em Ciências da Saúde (PPGCS) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela oportunidade de formação acadêmica e pela bolsa de estudos que tornou possível que eu realizasse a pós-graduação.

Agradeço ao meu Co-orientador Frederico Schmith Kremer, por ter aceitado entrar nesta jornada mesmo sabendo o quão difícil seria ter uma aluna que nunca trabalhou na área de Bioinformática. Devo agradacer a ele também pela amizade, paciência, conselhos, mas principalmente por sempre acreditar e confiar em mim.

Por fim, agradeço a minha orientadora Luciana Ávila, pela orientação, incentivo e confiança, mas principalmente pela empátia. Uma grande pessoa, sempre se mostrando humana, preocupada e fazendo muitas vezes um papel maior que da orientação. Agradeço por cada palavra de conforto nos momentos difíceis.

“O entusiasmo é a maior força da alma.
Conserva-o e nunca te faltará poder para
conseguires o que desejas.”

Napoleon Hill

RESUMO

Os *Lactobacillus acidophilus* pertencem ao grupo das Bactérias ácido lácticas, considerados probióticos, atuantes na promoção da saúde gastrointestinal, ocasionando inclusive inibição de microrganismos patogênicos. Este estudo, teve como objetivo prospectar, através de bioinformática, cepas de *L. acidophilus* que possuam potencial atividade probiótica. Inicialmente, foram acessados 13 genomas completos e 51 genomas rascunho de diferentes cepas de *L. acidophilus*, através do *National Center for Biotechnology Information* (NCBI). Uma análise pangênômica foi realizada visando identificar genes compartilhados entre as diferentes cepas, posteriormente foi utilizado o *BlastKOALA*, a fim de conhecer a quantidade de genes envolvidos na produção dos metabólitos. Para avaliar o potencial probiótico de cada cepa de *L. acidophilus*, utilizamos a ferramenta *iProbitics*, que facilita a triagem rápida de probióticos. Os resultados foram confrontados contra diferentes bancos de dados visando a identificação de características com potencial probiótico. Entre as 64 cepas testadas, destacamos que *L. acidophilus* NCTC 13721 e *L. acidophilus* DSM 20079, obtiveram maior quantidade genes envolvidos na produção metabólica de *quorum sensing*, ácido láctico, ácido acético e peróxido de hidrogênio, por meio do *BlastKOALA*. No *iProbitics*, a cepa NCTC 13721 e La-5 apresentaram 81,8% e 51,9% de probabilidade em serem *Lactobacillus* probióticos, respectivamente. Além disso, La-5 mostrou similaridade genética com as cepas DS20_1, LA-5, FSI4, APC2845, LA-G80-111, DS1_1A, LA1 e BCRC 14065, sendo necessárias mais avaliações sobre as mesmas. Por fim, a partir do uso de ferramentas de bioinformática conseguimos gerar um “*ranking*” de cepas probióticas de *L. acidophilus* que possam ser promissoras para estudos futuros.

Palavras-chave: Bioinformática; Probióticos; Bactérias ácido lácticas; Ácido láctico; Pangenoma;

ABSTRACT

Lactobacillus acidophilus belong to the group of lactic acid bacteria, considered probiotics, active in promoting gastrointestinal health, even causing inhibition of pathogenic microorganisms. This study aimed to prospect, through bioinformatics, strains of *L. acidophilus* that have potential probiotic activity. Initially, 13 complete genomes and 51 draft genomes of different strains of *L. acidophilus* were accessed through the National Center for Biotechnology Information (NCBI). A pangenomic analysis was carried out in order to identify genes shared between the different strains, later BlastKOALA was used in order to know the number of genes involved in the production of metabolites. To evaluate the probiotic potential of each strain of *L. acidophilus*, we used the iProbiotics tool, which facilitates the rapid screening of probiotics. The results were compared against different databases in order to identify characteristics with probiotic potential. Among the 64 strains tested, we highlight that *L. acidophilus* NCTC 13721 and *L. acidophilus* DSM 20079 obtained a greater number of genes involved in the metabolic production of quorum sensing, lactic acid, acetic acid and hydrogen peroxide, using BlastKOALA. In iProbiotics NCTC 13721 and La-5 showed 81.8%, 51.9% probability of being probiotic *Lactobacillus*, respectively. In addition, La-5 showed genetic similarity with strains DS20_1, LA-5, FSI4, APC2845, LA-G80-111, DS1_1A, LA1 and BCRC 14065, requiring further evaluation of these strains. Finally, from the use of bioinformatics tools, we managed to generate a ranking with probiotics strains of *L. acidophilus* that may be promising for future studies.

Keywords: Bioinformatics; Probiotics; Lactic acid bacteria; Lactic acid; Pangenome.

SUMÁRIO

1 INTRODUÇÃO E REVISÃO DE LITERATURA	12
1.1 Probióticos	12
1.3 <i>Lactobacillus acidophilus</i>	15
1.4 Bioinformática	16
2 OBJETIVOS	18
2.1 Objetivo geral	18
2.2 Objetivos específicos	18
3 BIBLIOGRAFIA	19
4 MANUSCRITO	23
5 CONSIDERAÇÕES FINAIS	47

1 INTRODUÇÃO E REVISÃO DE LITERATURA

1.1 Probióticos

Em meados de 1965, foi criada a primeira definição de probióticos, estes que na época foram denominadas substâncias produzidas por bactérias, que promovem o crescimento de outras bactérias (LILLY; STILLWELL, 1965; WIEËRS *et al.*, 2020). Na atualidade, os probióticos são definidos pela *The Food and Agriculture Organization (FAO) of the United Nations* como microrganismos vivos que, se administrados em quantidades adequadas, conferem benefícios à saúde do hospedeiro (WHO, 2001). Sua atuação na promoção de saúde, pode ocorrer através da estimulação da microbiota intestinal, imunidade e produção de metabólitos. Os principais metabólitos produzidos são: peróxido de hidrogênio, ácidos orgânicos e bacteriocinas, que podem atuar como antimicrobianos frente a bactérias e outros patógenos (YADAV *et al.*, 2022).

O trato gastrointestinal é a maior ligação do corpo com o meio externo e está em contínua comunicação com o sistema imunológico (LA FATA; WEBER; MOHAJERI, 2018), o qual é capaz de influenciar o microbioma intestinal, proporcionando uma composição benéfica da microbiota e inibindo o crescimento de bactérias específicas (LA FATA; WEBER; MOHAJERI, 2018). Os microrganismos que conseguem interagir de forma direta com bactérias patogênicas, são responsáveis pela regulação entre os patógenos e o sistema imunológico (LA FATA; WEBER; MOHAJERI, 2018). A homeostase intestinal ocorre quando os microrganismos mutualísticos, comensais e oportunistas se mantêm em equilíbrio, que é estabelecido somente quando o sistema imunológico consegue comunicação com a microbiota intestinal, sendo fundamental para manter a integridade da barreira intestinal (LA FATA; WEBER; MOHAJERI, 2018).

Desde os primeiros relatos sobre os probióticos, estes vêm sendo testados quanto às suas ações na profilaxia e tratamento de diversos distúrbios gastrointestinais. Porém, os mecanismos de ação dos probióticos não estão totalmente estabelecidos. Sendo assim, algumas pesquisas sugerem que eles atuam na inibição de microrganismos patogênicos, na prevenção da invasão de patógenos e na melhora da função da barreira intestinal (DALE *et al.*, 2019). Os probióticos promovem uma microbiota equilibrada, que consequentemente irá proporcionar um habitat intestinal favorável, apoiando ainda a saúde do trato digestivo e sistema imunológico, através da modulação e estimulação de respostas imunes locais e sistêmicas do hospedeiro (AGAMENON *et al.*, 2018).

São capazes de prevenir a patogenicidade, intervindo na sinalização entre microrganismos patogênicos, destruindo moléculas de detecção de *quorum sensing*, podendo também atuar na competição com patógenos por nutrientes e sítios de adesão na mucosa gastrointestinal. Além disso, possuem atividade antagonista, que acontece por meio da geração de metabólitos e bacteriocinas, com atividade antimicrobiana em oposição a microrganismos patogênicos (AGAMENON *et al.*, 2018).

Para serem considerados probióticos, os microrganismos necessitam de diferentes características, como: ser seguro para ingestão do hospedeiro, sobreviver ao trânsito pelo trato gastrointestinal e possuir benefícios para a saúde humana (TOURET; OLIVEIRA; LEMSADDEK, 2018). Estes microrganismos benéficos, precisam ser administrados no organismo através de alimentos e suplementos dietéticos, sendo importante frisar que seu efeito probiótico vai depender da dose ingerida (ZAKRZEWSKA *et al.*, 2022). Dentre os efeitos benéficos sugeridos destacam-se: o controle de infecções intestinais; o estímulo da motilidade intestinal, com consequente redução da constipação intestinal, promovendo melhor absorção de alguns nutrientes; melhora de características relacionadas à nutrição, fisiologia, metabolismo e imunidade, ocasionando bem estar ao hospedeiro (LIN *et al.*, 2019; GIAZZI, 2020).

Um dos mecanismos utilizados pelas bactérias probióticas é o *quorum sensing* (QS), que é um sistema de sinalização entre bactérias mediado por autoindutores (AI). Esse sistema atua na esporulação, competência, produção de antibióticos, e formação de biofilme (GUNARATNAM *et al.*, 2021). Conforme a densidade populacional aumenta, os AI se acumulam no ambiente e controlam a expressão gênica. O QS promove uma unidade coletiva, ou seja, um conjunto multicelular contra as células individuais (patógenos) (GUNARATNAM *et al.*, 2021).

1.3 *Lactobacillus acidophilus*

Dentre os diversos probióticos, as Bactérias ácido lácticas (BAL), formam um grupo de bactérias Gram-positivas, catalase-negativas, anaeróbias e não esporuladas, cujo principal produto da fermentação é o ácido láctico (GARBAZ, 2022). Estas bactérias são benéficas à saúde humana e animal, possuem aplicação na indústria de alimentos e medicamentos e são consideradas seguras (GRAS) (HUANG *et al.*, 2020). Os gêneros *Lactobacillus acidophilus* e *Bifidobacterium animalis*, estão ligados a benefícios para microbiota intestinal e são os mais

utilizadas na área industrial (CARNEIRO *et al.*, 2022). A fermentação produzida por essas bactérias têm a função de melhorar a textura e sabor dos alimentos (GALLINA, 2019). O crescimento das BAL ocasiona acidificação do material alimentício, desta forma conservando o produto (BULL *et al.*, 2013). As BAL produzem grande quantidade de ácidos e substâncias inibidoras de microrganismos patogênicos (GALLINA, 2019). Além disso, são acidófilas, característica que lhe confere sobrevivência e replicação em habitats com pH baixo, (GARBAZ, 2022) são tolerantes a bile e a ácidos, não patogênicas e capazes de produzir substâncias antimicrobianas (REUBEN *et al.*, 2019). O peróxido de hidrogênio e ácidos orgânicos, destacando o ácido láctico, produzem o efeito inibitório microbiano (GASPAR *et al.*, 2018). Algumas das características das Bactérias ácido lácticas, grupo ao qual pertencem os *Lactobacillus*, podem ser visualizadas na Tabela 1.

Tabela 1. Na Tabela 1, é possível observar diferentes benefícios dos probióticos, dos gêneros amplamente utilizados *Bifidobacterium* spp. e *Lactobacillus* spp. , bem como a metodologia utilizada no estudo.

Autor/Ano	Probiótico	Metodologia	Benefícios
Slaterry; Cotter; O'Toole, 2019	<i>Lactobacillus</i> <i>plantarum</i> CIDCA 83114	<i>In vitro/In vitro</i>	- Efeito protetor contra patógenos em células humanas cultivadas. - Diminuição de diarreia e mortalidade em hamster contra <i>Clostridium difficile</i> .
Wang <i>et al.</i> , 2019	<i>Bifidobacterium</i> <i>longum</i> 1714	Estudo randomizado, duplo-cego controlado por placebo	-Redução nos comportamentos relacionados ao estresse. - Melhora em respostas ao estresse e função cognitiva.
Yan <i>et al.</i> , 2019	<i>Lactobacillus</i> <i>acidophilus</i> KLDS1.1003 e <i>Lactobacillus</i> <i>acidophilus</i> KLDS1.0901	<i>In vivo</i>	- Melhora na função da barreira epitelial. - Diminuição de citocinas inflamatórias. -Regulação da expressão de genes relacionados ao metabolismo de glicose e lipídeos.

Baoqing <i>et al.</i> , 2022	<i>Bifidobacterium lactis</i> Pro-Bio-M8	Ensaio randomizado, duplo-cego controlado por placebo	-Modulação da microbiota intestinal. -Coadministração do probiótico auxiliou na melhora do tratamento clínico da doença arterial coronariana.
Hu <i>et al.</i> ,	<i>Lactobacillus casei</i> LH23	<i>In vitro</i>	- Lisados do probiótico suprimem a proliferação de células de câncer cervical. -Papel inibitório dos lisados do probiótico, na expressão de oncogenes do vírus HPV em células de câncer cervical.

Os *L. acidophilus* pertencem as BAL, podem ser encontrados na vagina, cavidade oral e trato gastrointestinal humano, é considerado um probiótico seguro, atuante na promoção da saúde gastrointestinal, sendo grandes produtores de ácido láctico (HUANG *et al.*, 2022). É importante mencionar que, *L. acidophilus* é uma das espécies mais relevantes comercialmente, e possui diversas cepas que são utilizadas como cultura probiótica nas indústrias de laticínios, pois dispõem de um grande histórico de segurança e benefícios para a saúde (Figura 1) (GOH; BARRANGOU; KLAENHAMMER, 2021).

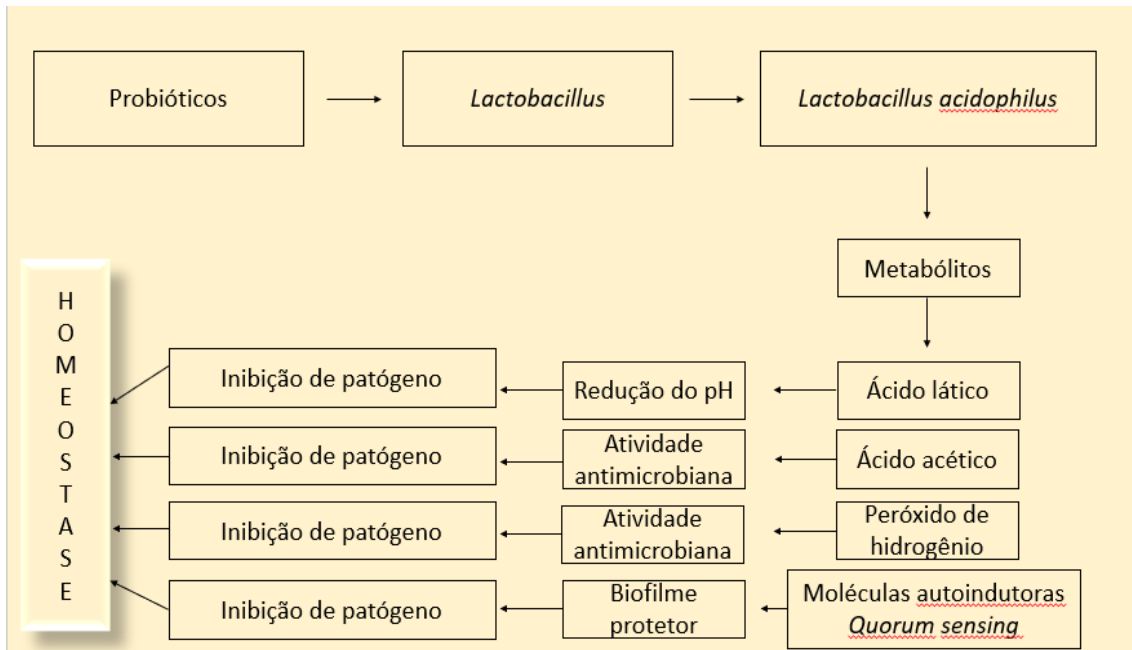


Figura 1- *Lactobacillus acidophilus* e metabólitos associados que promovem benefícios ao hospedeiro.

A sobrevivência de *L. acidophilus* no habitat intestinal, necessita de mecanismos de adaptação e condições extremas, presença de ácidos biliares, compostos antimicrobianos, utilização de vários carboidratos e fontes de nutrientes no intestino (GOH; BARRANGOU; KLAENHAMMER, 2021). Além disso, efeitos benéficos desta espécie vêm sendo relatados, destacando-se a redução do pH no lúmen, crescimento de células epiteliais gerando aumento do fluxo sanguíneo e modificação da motilidade intestinal, auxílio na absorção de minerais e água e elevação na produção de muco (OHASHI; USHIDA, 2009; CADORE *et al.*, 2021).

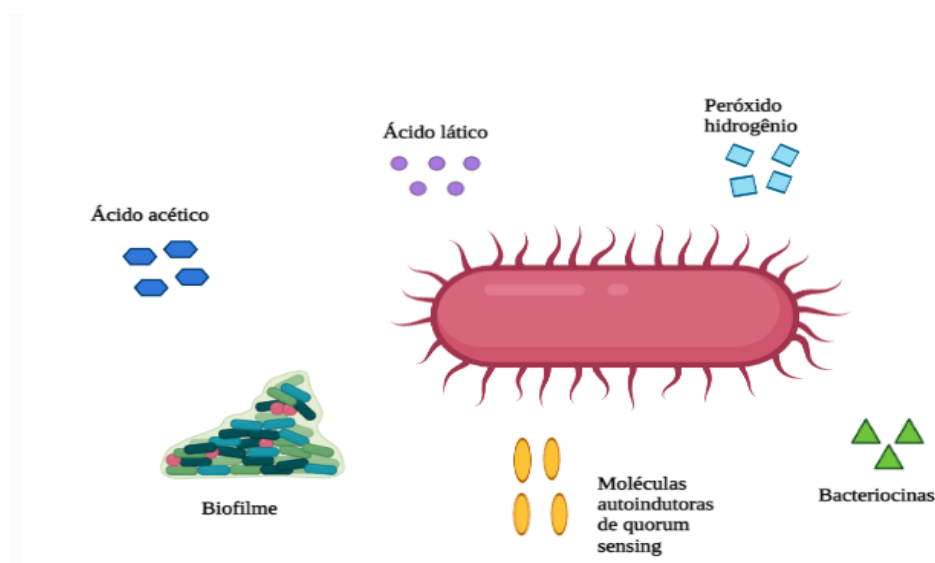


Figura 2- Uma representação de metabólitos secretados de bactérias, que possuem atividade inibitória contra patógenos. A formação de biofilme, pode gerar uma proteção, sendo capaz de impedir o patógeno de

aderir-se ao sítio de adesão. Ácido láctico, peróxido de hidrogênio, ácido acético, bacteriocinas, são capazes de servir como moléculas de detecção de *quorum* e/ou modular diretamente a composição do microbioma

Em relação aos outros probióticos, a espécie se sobressai, já que consegue sobreviver em meios mais ácidos do que os demais e tem como característica a produção de metabólitos antimicrobianos (Figura 2) (MENEZES *et al.*, 2022). Cheng *et al.*, (2022) aponta que os *L. acidophilus*, conseguem sobreviver em pH baixo, fato que faz com que eles consigam atingir o intestino delgado, possuem também efeitos benéficos como manutenção da homeostase imunológica, atividade antiviral e antibiótica, bem como, absorção de eletrólitos.

1.4 Bioinformática como ferramenta de prospecção de organismos probióticos

O estudo das bactérias sofreu grande mudança ao passar dos anos, tendo ganhado destaque no campo tecnológico (KOBAS; FENTON; SHEPPARD, 2021). Devido a isso, atualmente entendemos mais sobre o metabolismo, interações dos microrganismos com o hospedeiro e sobre o contexto ambiental (KOBAS; FENTON; SHEPPARD, 2021). As técnicas de sequenciamento de nova geração (NGS), possibilitaram que os microrganismos fossem estudados de uma forma mais ampla, visto que proporcionam uma caracterização otimizada dos genomas e identificação taxonômica robusta (CAO *et al.*, 2019). É importante ressaltar ainda que, existem poucas informações sobre probióticos promissores e sua capacidade metabólica associada. As técnicas de NGS são utilizadas nas indústrias de laticínios, a fim de proporcionar uma melhor caracterização das cepas, o que consequentemente possibilita aumentar o potencial probiótico (CAO *et al.*, 2019; SHAKIRA *et al.*, 2020).

A análise do pangenoma é caracterizada pelo estudo do conjunto dos diferentes genes que constituem o repertório gênico de um grupo de organismos, geralmente bacterianos. Ademais, o pangenoma pode ser dividido em diferentes frações, dentre elas o genoma núcleo, também conhecido como *core genes*, constituído por genes que estão presentes em todos os organismos analisados; o genoma acessório, chamado também de *shell genes* e *cloud genes*, é integrado por genes que devem estar presentes ao menos em dois organismos (PARISE, 2016). Os genes são agrupados em conjuntos que são denominados *clusters*, conforme a similaridade do alinhamento destes, os *clusters* são constituídos por grupos de genes parálogos e ortólogos, sendo o último genes originados a partir da especiação de um gene ancestral comum (PARISE, 2016). Dois parâmetros genotípicos são essenciais para atribuir

uma cepa a uma espécie: grau de hibridização DNA-DNA (DDH) acima de 70% e a similaridade de sequência acima de 98,7% no gene 16S rRNA (WAYNE *et al.*, 1987; INGLIN; MEILE; STEVENS, 2018). Uma pesquisa utilizando 7 diferentes cepas de *L. acidophilus*, obteve como resultado um pangenoma de 2.254 de genes, um genoma central com 1.726 genes e 200 genes acessórios (SON *et al.*, 2021). Ainda, a análise pangenômica de 105 genomas de *Lactobacillus crispatus* identificou um pangenoma com 12.114; 465 genes conservados; 4275 genes únicos (PAN; HIDALGO-CANTABRANA; BARRANGO, 2020).

Portanto, a bioinformática é uma área de conhecimento interdisciplinar que utiliza ferramentas como ciência da computação, matemática e biologia, a fim de promover um maior entendimento sobre os dados biológicos. Desta maneira, beneficiando as pesquisas biomédicas em diversos aspectos de modo a compreender a relação entre os genes e doenças, interpretando o mecanismo destas, sendo assim, melhorando diagnóstico, prevenção e tratamento (LIANG *et al.*, 2019). Diante disso, torna-se importante a realização de análises *in silico* de *Lactobacillus acidophilus* a fim de ampliar os conhecimentos sobre genes envolvidos em processos metabólicos que possam estar relacionados à ação probiótica, servindo de base para estudos *in vitro* e *in vivo*.

2 OBJETIVOS

2.1 Objetivo geral

Prospectar, através de Bioinformática, cepas de *Lactobacillus acidophilus* que possuem potencial atividade probiótica.

2.2 Objetivos específicos

- Identificar através de análise pangenômica grupos ortólogos presentes em diferentes cepas de *Lactobacillus acidophilus*.
- Identificar através de bioinformática a quantidade genes envolvidos na produção de ácido láctico, ácido acético, peróxido de hidrogênio e *quorum sensing* relacionados à atividade probiótica.
- Identificar através de similaridade genômica, cepas com potencial probiótico.

3 BIBLIOGRAFIA

AGAMENNONE, V. *et al.* A practical guide for probiotics applied to the case of antibiotic-associated diarrhea in The Netherlands. **BMC Gastroenterol** 18, 103 (2018). <https://doi.org/10.1186/s12876-018-0831-x>.

BULL, M *et al.* The life history of *Lactobacillus acidophilus* as a probiotic: a tale of revisionary taxonomy, misidentification and commercial success. **FEMS Microbiol Lett.** 2013 Dec;349(2):77-87. doi: 10.1111/1574-6968.12293.

CADORE, P.S. *et al.* Protective effect of the probiotic *Lactobacillus acidophilus* ATCC 4356 in BALB/c mice infected with *Toxocara canis*. **Rev Inst Med Trop Sao Paulo.** 2021 Jan 29;63:e9. doi: 10.1590/S1678-9946202163009.

CAO, Y. *et al.* A Review on the Applications of Next Generation Sequencing Technologies as Applied to Food-Related Microbiome Studies. **Front Microbiol.** 2017 Sep 21;8:1829. doi: 10.3389/fmicb.2017.01829.

CARNEIRO, M. DA S. *et al.* Processing of soy beverages obtained from the grain, flour and powder extract and fermented by probiotics. **Food Science and Technology**, v. 42, n. Food Sci. Technol, 2022 42, 2022.

CARVALHO, C. *et al.* The Relevance of In Silico, In Vitro and Non-human Primate Based Approaches to Clinical Research on Major Depressive Disorder. **Altern Lab Anim.** 2019 Jul-Sep;47(3-4):128-139. doi: 10.1177/0261192919885578.

CHENG, H. *et al.* A Systematic Review and Meta-Analysis: *Lactobacillus acidophilus* for Treating Acute Gastroenteritis in Children. **Nutrients.** 2022 Feb 6;14(3):682. doi: 10.3390/nu14030682.

CRISTOFORI, F *et al.* Anti-Inflammatory and Immunomodulatory Effects of Probiotics in Gut Inflammation: A Door to the Body. **Front Immunol.** 2021 Feb 26;12:578386. doi: 10.3389/fimmu.2021.578386.

DALE, H.F, *et al.* Probiotics in Irritable Bowel Syndrome: An Up-to-Date Systematic Review. **Nutrients.** 2019 Sep 2;11(9):2048. doi: 10.3390/nu11092048.

FAO/WHO. World Health Organization. Joint FAO/WHO Consultation on Evaluation of Health and Nutritional Properties of Probiotics in Food Including Powder Milk with Live Lactic Acid Bacteria, 2001.

GALLINA, D. A. *et al.* Effect of modified chitooligosaccharides on the physicochemical and microbiological characteristics of a probiotic fermented beverage. **Revista do Instituto de Laticínios Cândido Tostes**, v. 74, n. 4, p. 240-251, 2019.

GARBACZ K. Anticancer activity of lactic acid bacteria. **Semin Cancer Biol.** 2022 Nov;86(Pt3):356-366. doi: 10.1016/j.semcancer.2021.12.013.

GASPAR, C. *et al.* Bacteriocin production of the probiotic *Lactobacillus acidophilus* KS400. **AMB Express.** 2018 Sep 27;8(1):153. doi: 10.1186/s13568-018-0679-z.

GIAZZI, A. *et al.* Propriedades tecnológicas das bactérias ácido lácticas isoladas na região norte do Paraná. **Brazilian Journal of Development**, v. 6, n. 4, p. 18861-18877, 2020.

GOH, Y.J.; BARRANGOU, R.; KLAENHAMMER, T.R. *In Vivo* Transcriptome of *Lactobacillus acidophilus* and Colonization Impact on Murine Host Intestinal Gene Expression. **mBio.** 2021 Jan 26;12(1):e03399-20. doi: 10.1128/mBio.03399-20.

HU, S. *et al.* *Lacticaseibacillus casei* LH23 Suppressed HPV Gene Expression and Inhibited Cervical Cancer Cells. **Probiotics Antimicrob Proteins.** 2021 Oct 2. doi: 10.1007/s12602-021-09848-7.

HUANG, Z. *et al.* Different Effects of Different *Lactobacillus acidophilus* Strains on DSS-Induced Colitis. **Int J Mol Sci.** 2022 Nov 27;23(23):14841. doi: 10.3390/ijms232314841. PMID: 36499169.

HUANG, H.C. *et al.* Lactic Acid Bacteria and Lactic Acid for Skin Health and Melanogenesis Inhibition. **Curr Pharm Biotechnol.** 2020;21(7):566-577. doi: 10.2174/1389201021666200109104701.

INGLIN, R.C.; MEILE, L.; STEVENS, M.J.A. Clustering of Pan- and Core-genome of *Lactobacillus* provides Novel Evolutionary Insights for Differentiation. **BMC Genomics** 19, 284 (2018). <https://doi.org/10.1186/s12864-018-4601-5>.

KIM, S.K. *et al.* Role of Probiotics in Human Gut Microbiome-Associated Diseases. **J Microbiol Biotechnol.** 2019 Sep 28;29(9):1335-1340. doi: 10.4014/jmb.1906.06064.

KOBRAS, C.M.; FENTON, A.K.; SHEPPARD, S.K. Next-generation microbiology: from comparative genomics to gene function. **Genome Biol** 22, 123 (2021). <https://doi.org/10.1186/s13059-021-02344-9>.

- LA FATA G; WEBER P; MOHAJERI MH. Probiotics and the Gut Immune System: Indirect Regulation. **Probiotics Antimicrob Proteins**. 2018 Mar;10(1):11-21. doi: 10.1007/s12602-017-9322-6.
- LIANG, X. *et al.* Molecular Computing and Bioinformatics. **Molecules**. 26;24(13):2358.2019. doi: 10.3390/molecules24132358.
- LILLY, D.M; STILLWELL, R H (1965). Probiotics: growth-promoting factors produced by microorganisms. **Science** 147,747-748. doi: 10.1126/science.147.3659.747.
- LIN, T. L. *et al.* Investiture of next generation probiotics on amelioration of diseases- Strains do matter. **Medicine in Microecology**. v. 1-2,2019, 100002, ISSN 2590-0978, <https://doi.org/10.1016/j.medmic.2019.100002>.
- LIU Y; TRAN DQ; RHOADS JM. Probiotics in Disease Prevention and Treatment. **J Clin Pharmacol**. 2018 Oct;58 Suppl 10(Suppl 10):S164-S179. doi: 10.1002/jcph
- MENEZES, M. U. F. O. *et al.* Viability of *Lactobacillus acidophilus* in whole goat milk yogurt during fermentation and storage stages: a predictive modeling study. **Food Science and Technology**, v. 42, n. Food Sci. Technol, 2022 42, 2022.
- OHASHI Y, USHIDA K. Health-beneficial effects of probiotics: its mode of action. **Anim Sci J**. 2009;80:361-71.
- PAN, M.; HIDALGO-CANTABRANA, C.; BARRANGOUR. Host and body site-specific adaptation of *Lactobacillus crispatus* genomes. **NAR Genom Bioinform**. 2020 Jan 24;2(1):lqaa001. doi: 10.1093/nargab/lqaa001.
- PARISE, Douglas. **Genômica comparativa entre os biovares Equi e Ovis de Corynebacterium pseudotuberculosis isolados no México**. 2016. Dissertação (Mestrado em Bioinformática) - Programa Interunidades de Pós-Graduação em Bioinformática, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte,2016.
- REUBEN R.C. *et al.* Characterization and evaluation of lactic acid bacteria from indigenous raw milk for potential probiotic properties. **J Dairy Sci**. 2020 Feb;103(2):1223-1237. doi: 10.3168/jds.2019-17092.
- SHAKIRA G. *et al.* "Role of Next Generation Sequencing (NGS) in Searching for Promising Lactic Acid Strains Used as Animal Probiotic". **Acta Scientific Nutritional Health** 4.4 (2020): 179-183.

SLATTERY, C; COTTER, P.D; O'TOOLE P.W. Analysis of Health Benefits Conferred by *Lactobacillus* Species from Kefir. **Nutrients**. 2019 Jun 1;11(6):1252. doi: 10.3390/nu11061252.

SON, S. *et al.* Complete genome sequencing and comparative genomic analysis of *Lactobacillus acidophilus* C5 as a potential canine probiotics. **J Anim Sci Technol**. 2021 Nov;63(6):1411-1422. doi: 10.5187/jast.2021.e126.

SUN, B. *et al.* Bifidobacterium lactis Probio-M8 Adjuvant Treatment Confers Added Benefits to Patients with Coronary Artery Disease via Target Modulation of the Gut-Heart/-Brain Axes. **mSystems**. 2022 Apr 26;7(2):e0010022. doi: 10.1128/msystems.00100-22.

TOURET T, OLIVEIRA M, SEMEDO-LEMSADDEK T. Putative probiotic lactic acid bacteria isolated from sauerkraut fermentations. **PLoS One**. 2018 Sep7;13(9):e0203501. doi:10.1371/journal.pone.0203501.

WAYNE, L.G. *et al.* Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. **Int J Syst Bacteriol** 1987;37:463–464. doi:https://doi.org/10.1099/00207713-37-4-463.

WIEËRS, G. *et al.* How Probiotics Affect the Microbiota. **Front Cell Infect Microbiol**. 2020 Jan 15;9:454. doi: 10.3389/fcimb.2019.00454.

WANG, H.; BRAUN, C.; MURPHY, E.F.; ENCK, P. Bifidobacterium longum 1714™ Strain Modulates Brain Activity of Healthy Volunteers During Social Stress. **Am J Gastroenterol**. 2019 Jul;114(7):1152-1162. doi: 10.14309/ajg.0000000000000203.

YADAV, M.K. *et al.* Probiotics, prebiotics and synbiotics: Safe options for next-generation therapeutics. **Appl Microbiol Biotechnol**. 2022 Jan;106(2):505-521. doi: 10.1007/s00253-021-11646-8.

YAN, F. *et al.* *Lactobacillus acidophilus* alleviates type 2 diabetes by regulating hepatic glucose, lipid metabolism and gut microbiota in mice. **Food Funct**. 2019 Sep 1;10(9):5804-5815. doi: 10.1039/c9fo01062a

ZAKRZEWSKA, Z. *et al.* Prebiotics, Probiotics, and Postbiotics in the Prevention and Treatment of Anemia. **Microorganisms**. 2022 Jun 30;10(7):1330. doi: 10.3390/microorganisms10071330

4 MANUSCRITO

Manuscrito a ser submetido à revista *Brazilian Journal of Microbiology*

Qualis: A2

Fator de impacto: 2.214

SCImago Journal Rank (SJR): 0,643

***In silico* prospection of *Lactobacillus acidophilus* strains with potential probiotic activity**

Rafaella Sinnott Dias¹, Frederico Schmith Kremer², Luciana Farias da Costa de Avila¹

¹ Universidade Federal do Rio Grande – FURG, Faculty of Medicine, Academic Area of the University Hospital, Post-Graduate Program in Health Sciences – Parasitology Laboratory, Rio Grande, RS, Brasil" Corresponding author- email: lucostaavila@hotmail.com

² Bioinformatics Laboratory, Technological Development Center - Federal University of Pelotas. Capão do Leão, Rio Grande do Sul, Brazil.

Abstract

Lactic acid bacteria (LAB) are widely studied fermentative microorganisms, as they play different roles in biotechnological processes, mainly in the food and pharmaceutical industries. Among the (BAL), *Lactobacillus acidophilus* is a species that deserves to be highlighted for being used both in the prophylaxis and in the treatment of pathologies. Most of the metabolites generated by the species are linked to the inhibition of pathogens, a fact that makes it so promising. In this study, we used different bioinformatics strategies to list promising strains of *L. acidophilus* with potential probiotic activity. Among the 64 strains tested, NCTC 13721, DSM 20079, proved to be the most promising for prospecting new probiotic *Lactobacillus*, since they showed a greater number of genes involved in the production of lactic acid, acetic acid, hydrogen peroxide and quorum sensing. The NCTC 13721 and La-5 strains showed greater ability to be *Lactobacillus* with probiotic capacity, showing 84.8% and 51.9%, respectively. In addition, eight strains with characteristics similar to La-5 were listed: S20_1, LA-5, FSI4, APC2845, LA-G80-111, DS1_1A, LA1 and BCRC 14065. Corroborating with the findings in iProbiotics and BlastKoala, a ranking with the most promising strains to be used for future *in vitro* studies.

Keywords: Bioinformatics; Probiotics; lactic acid bacteria; *Lactobacillus*; Lactic acid; Pangenome;

Problem Session: Molecular Biology of Bacteria, Fungi and Viruses.

Introduction

Probiotics are defined as live microorganisms which, if administered in adequate amounts, confer a health benefit on the host (Who 2001). Among the various probiotics, *Lactobacillus*, comprising 25 genera (including 23 new ones), represent a group of bacteria with heterogeneous characteristics, having the following classification: phylum *Firmicutes*, class *Bacilli*, order *Lactobacillales*, family *Lactobacillaceae* (Zheng et al. 2020). They are non-spore-forming, catalase-negative, Gram-positive, and facultatively anaerobic rods (Goldstein; Tyrrell and Citron 2015). It is also important to point out that *Lactobacillus* spp. are important producers of lactic acid, which gives them the name lactic acid bacteria (LAB), being present in various foods such as meat and fermented products (Ojha; Shah and Mishra 2021). Among the many protective effects of *Lactobacillus* spp. there is the modulation of the intestinal microbiota and the immune system, reinforcement of the intestinal barrier and regulation of crucial pathways in epithelial cells (Xiao et al. 2021).

L. acidophilus are currently subdivided into several strains with different functions and probiotic properties (Gao et al. 2022). This species contains important probiotics, as it acts in several functions related to the health of the host, mainly in the inhibition of pathogenic microorganisms, regulation of the intestinal epithelial barrier and anti-inflammatory effect, increasing its use in the food and pharmaceutical area (Li et al. al. 2020). In addition, *L. acidophilus* stands out for having characteristics such as resistance to bile salts, low pH, good ability to adhere to human colon cells in cell culture, regulation of host immune responses, showing promise in the prophylaxis and treatment of infections. Among its functional mechanisms, current studies have shown that *L. acidophilus* regulates the intestinal microbiota by decreasing pH and producing metabolites (Gao et al. 2022).

Bioinformatics is an interdisciplinary area of knowledge that uses tools such as computer science, mathematics and biology, in order to promote a greater understanding of biological data, benefiting biomedical research in several aspects in order to understand the relationship between genes and the system stimulus (Liang et al. 2019). Studies with pangenomics are becoming increasingly frequent, as genetic analysis and comparison can be useful to explore and characterize a shared pattern among microorganisms, providing a better understanding of the function and evolution of genomes (Huang et al. 2021). It is also possible that certain genomic characteristics are evaluated in order to know and characterize a species with better

performance for the production of a certain protein or metabolite, for example (Huang et al. 2021).

The objective of this work was to prospect, through bioinformatics, strains of *L. acidophilus* that have potential probiotic activity through pangenomic analysis and predictive models. In addition, we identified through bioinformatics the number of genes involved in the production of lactic acid, acetic acid, hydrogen peroxide and quorum sensing related to probiotic activity.

Materials and methods

Data acquisition

Complete genome sequences of *L. acidophilus* strains were accessed from the National Center for Biotechnology Information (NCBI). The total of genomes found was 64 *L. acidophilus*, being 13 complete genomes and 51 draft genomes. All those that did not have the same taxonomy as *L. acidophilus* (eg: phages) or did not have the sequence available in GenBank and FASTA format were excluded.

Pangenome analysis

After selecting the *Lactobacillus* strains, the Roary v3.13 software – native to Linux- (Carpi et al. 2021) was used, which receives GFF3 files (General Feature Formats version 3) as input. It was established as part of the core gene, the genes that are present in at least 95% of the genomes and the minimum similarity between two genes must be 70% for them to belong to the same orthologous group (Sitto and Battistuzzi 2020). The mentioned software provides several files with statistics of genes shared by a large part of the lineage or throughout (soft core and core genes) and also by some genomes (accessories, subdivided into cloud and shell genes) (Sitto and Battistuzzi 2020).

Identification of genes associated with the production of metabolites and probiotic activity

The genes involved in the production of lactic acid were obtained from the KEGG database, which has the function of linking biological functions. Subsequently, the BlastKOALA

program was applied, which is a tool used for annotation and can identify proteins involved in signal transduction, catabolism, transport, biosynthesis, glycan metabolism, among other metabolic pathways available from the KEGG database (Fuentes et al. 2021). After the metabolic annotation of the pangenome, the identified results were compared with the EC codes (Enzyme Commission Number) described in the KEGG database as associated with the production of lactic acid, acetic acid, hydrogen peroxide and quorum sensing. The number of genes identified for each biological process in each strain was then scaled, based on the greatest number of gene occurrences for each biological process across all strains. Based on this value, an average was calculated to reflect an empirical score.

Probiotic capacity analysis

For the analysis of the potential probiotic capacity of the different strains of *L. acidophilus*, the tool called iProbiotics was used, which facilitates the rapid screening of probiotics, based on the prediction of probiotic activity *in silico*, from the genome, which was obtained in FASTA format. iProbiotics has three different parameters: predictor of probiotic and non-probiotic strains (model one); predictor of *Lactobacillus probiotics*, *Bifidobacterium probiotics* and other probiotics (model two); and predictor of probiotic *Lactobacillus* and non-probiotic *Lactobacillus* (model three), models one and three being used in this work. This tool uses characteristics to define probiotic capacity, such as: composition of oligonucleotides, since it plays the role of a molecular marker, and genes related to probiotic function, such as: adsorption gene, competitiveness gene, gene related to growth rate, hydrolase gene of bile salts and gene related to retention (Sun et al. 2022).

Similarity analysis between genomes

Similarity analysis between genomes was performed using the FastANI tool, which is a method that estimates the Average Nucleotide Identity (ANI) through sequence comparison without alignment (Jain et al. 2018). In all, 4096 ANI comparisons were produced, since the analysis is carried out in an “all against all” (all-vs-all) way. Briefly, the ANI technique allows estimating a global similarity between two genomes Figure 3, also serving as an indicator for taxonomic classification of genera and species (Ciufo et al. 2018).

Results

In Table 1, we can see the 64 strains of *L. acidophilus*, their NCBI accession code and completion level. Among the 64 strains used, 13 have complete genomes and 51 draft genomes.

Table 1- Identification and access codes of the *L. acidophilus* strains used in the present study.

Strain	Code NCBI	Completion level
NCFM	GCA_000011985.1	Complete
FSI4	GCA_000934625.1	Complete
LA1	GCA_002286215.1	Complete
YT1	GCA_003952845.1	Complete
LA-G80-111	GCA_013342945.1	Complete
NC55	GCA_020883435.1	Complete
5460	GCA_021432145.1	Complete
La-14	GCF_000389675.2	Complete
ATCC 53544	GCF_002224305.1	Complete
DSM 20079	GCF_003047065.1	Complete
HN017	GCF_024397395.1	Complete
LA-2	GCF_024665075.1	Complete
LA-5	GCF_024665555.1	Complete
NCTC13721	GCA_900452495.1	Draft
KLDS 1.0901	GCF_001868765.1	Draft

LA1063	GCF_017009725.1	Draft
LMG P-21904	GCF_002914945.1	Draft
BCRC 14065	GCF_017009515.1	Draft
BCRC 17008	GCF_017009595.1	Draft
BCRC 17481	GCA_017009655.1	Draft
NBRC 13951	GCF_001591845.1	Draft
La-5	GCA_017009715.1	Draft
CIRM-BIA 442	GCF_000442865.1	Draft
ATCC 4356	GCA_000786395.1	Draft
BCRC 16092	GCA_017009575.1	Draft
BCRC 16099	GCA_017009605.1	Draft
NBIMCC 8242 (180)	GCF_021229035.1	Draft
DSM 20242	GCA_000442825.1	Draft
QAULAN51	GCA_022509485.1	Draft
s-4	GCF_013867555.1	Draft
BCRC 14079	GCF_017009475.1	Draft
CIRM-BIA 445	GCA_000469765.1	Draft
s-13	GCF_013867605.1	Draft
BCRC 80064	GCF_017009695.1	Draft
L3_101_000G1_dasL3_101_000G1_metabat.metabat.48	GCA_018367455.1	Draft
BCRC 17486	GCF_017009585.1	Draft
PNW3	GCA_004348805.1	Draft
MGYG-HGUT-02379	GCF_902386525.1	Draft

DSM 9126	GCA_000469745.1	Draft
LA_AVK2	GCF_009741835.1	Draft
LA_AVK1	GCF_009742735.1	Draft
BCRC 12255	GCF_017009485.1	Draft
DS9_1A	GCF_003061925.1	Draft
DS5_1A	GCF_003061985.1	Draft
BIO6307	GCF_008868625.1	Draft
DSM 20079	GCA_001433895.1	Draft
DS24_1	GCA_003053135.1	Draft
DS8_1A	GCA_003061945.1	Draft
DS10_1A	GCA_003053245.1	Draft
CIP 76.13	GCF_000469705.1	Draft
DS11_1A	GCF_003062025.1	Draft
UBLA-34	GCF_003641085.1	Draft
DS20_1	GCA_003061885.1	Draft
DS1_1A	GCA_003062045.1	Draft
DS13_1B	GCF_003061905.1	Draft
ATCC 4796	GCA_000159715.1	Draft
DS13_1A	GCF_003061965.1	Draft
APC2845	GCA_017695935.1	Draft
LA-G80	GCF_018252545.1	Draft
PB2021-BA04	GCA_023093425.1	Draft
P2	GCF_002406675.1	Draft

WG-LB-IV	GCF_001639165.1	Draft
DS2_1A	GCA_003062005.1	Draft
CFH	GCF_000497795.1	Draft

Table 2 shows the numbers of genes identified by the Roary tool from the pangenome, showing the core genome, pangenome and accessory genome. A representation of the size distribution of the pangenome and core genome is shown in Figure 1.

Table 2. Number of genes present in the pangenome (total genes), core genome (core genes and soft core genes), and accessory genome (shell genes and cloud genes), as calculated by the Roary tool. The core genome is composed of conserved genes in at least 95% of the analyzed strains.

Gene pool	Abundance	Amount
Core genes	(95% <= strains <= 100%)	1506
Soft core genes	(94% <= strains < 95%)	0
Shell genes	(15% <= strains < 94%)	485
Cloud genes	(0% <= strains < 15%)	2158
Total genes	(0% <= strains <= 100%)	4149

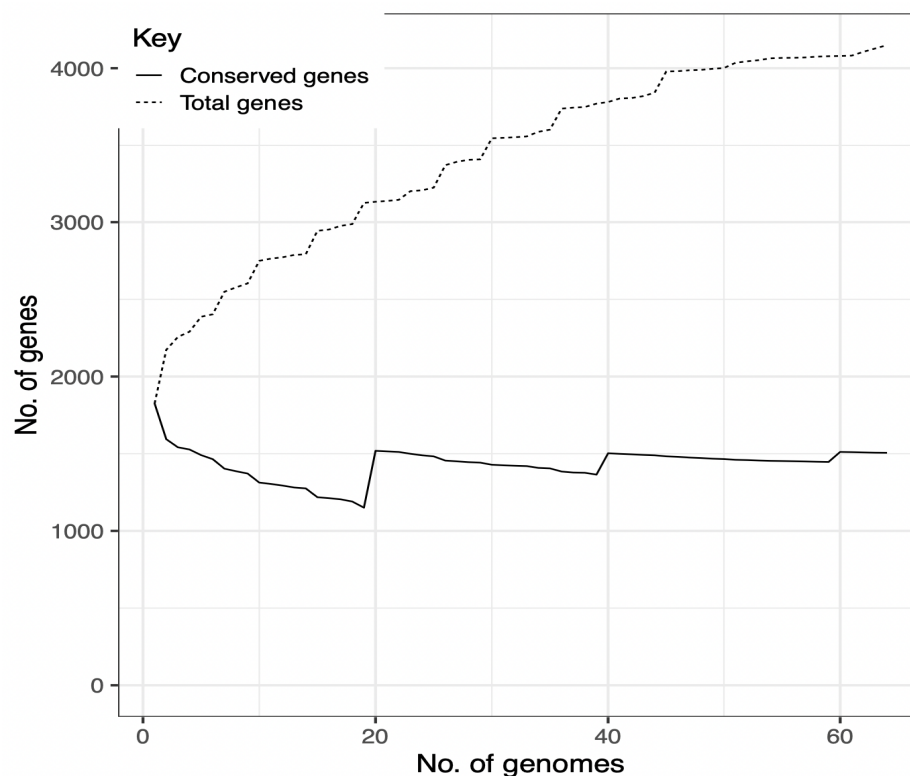


Figure 1- Graph showing the size of the pangenome and conserved genes (core genome) for different iterations from Roary analysis.

From the visualization of Table 3, we noticed genes involved in the production of metabolites such as acetic acid, hydrogen peroxide, lactic acid and quorum sensing, in addition to the score generated by these genes present in the different strains. Among the 64 strains of *L. acidophilus* tested, NCTC 13721 stood out, as it presented a greater number of genes involved in the quorum sensing process than the others.

Table 3. Number of genes involved in the metabolic production of different strains of *L. acidophilus* identified from the analysis with BLASTKOALA.

Strain	Genome	Production metabolic				Score
		Acid acetic	Hydrogen peroxide	Acid lactic	Quorum Sensing	
NCTC 13721	GCA_900452495.1	21	21	10	38	0.91
DSM 20079	GCF_003047065.1	22	22	9	24	0.77

APC2845	GCA_017695935.1	21	21	7	26	0.71
DSM 20079	GCA_001433895.1	20	20	7	29	0.71
ATCC 4796	GCA_000159715.1	20	20	7	28	0.70
NCFM	GCA_000011985.1	20	20	7	27	0.70
DSM 20242	GCA_000442825.1	20	20	7	27	0.70
DSM 9126	GCA_000469745.1	20	20	7	27	0.70
DS24_1	GCA_003053135.1	20	20	7	26	0.70
DS10_1A	GCA_003053245.1	20	20	7	26	0.70
DS20_1	GCA_003061885.1	20	20	7	26	0.70
DS8_1A	GCA_003061945.1	20	20	7	26	0.70
DS2_1A	GCA_003062005.1	20	20	7	26	0.70
DS1_1A	GCA_003062045.1	20	20	7	26	0.70
CIRM-BIA 445	GCA_000469765.1	21	21	7	28	0.69
YT1	GCA_003952845.1	20	20	7	25	0.69
La-5	GCA_017009715.1	19	19	7	27	0.69
LA1	GCA_002286215.1	19	19	7	26	0.68
LA-G80-111	GCA_013342945.1	19	19	7	26	0.68
BCRC 16092	GCA_017009575.1	19	19	7	26	0.68
BCRC 16099	GCA_017009605.1	19	19	7	26	0.68
BCRC 17481	GCA_017009655.1	19	19	7	26	0.68

L3_101_000 G1_dasL3_10 1_000G1_me tabat.metabat. 48	GCA_018367455.1	19	19	7	26	0.68
BIO6307	GCF_008868625.1	19	19	7	26	0.68
LA_AVK2	GCF_009741835.1	19	19	7	26	0.68
LA1063	GCF_017009725.1	19	19	7	26	0.68
ATCC 4356	GCA_000786395.1	19	19	7	25	0.68
FSI4	GCA_000934625.1	19	19	7	25	0.68
QAULAN51	GCA_022509485.1	19	19	7	25	0.68
PB2021-BA0 4	GCA_023093425.1	19	19	7	25	0.68
La-14	GCF_000389675.2	19	19	7	25	0.68
CIRM-BIA 442	GCF_000442865.1	19	19	7	25	0.68
CIP 76.13	GCF_000469705.1	19	19	7	25	0.68
CFH	GCF_000497795.1	19	19	7	25	0.68
WG-LB-IV	GCF_001639165.1	19	19	7	25	0.68
KLDS 1.0901	GCF_001868765.1	19	19	7	25	0.68
ATCC 53544	GCF_002224305.1	19	19	7	25	0.68
P2	GCF_002406675.1	19	19	7	25	0.68
BA05	GCF_002914945.1	19	19	7	25	0.68

DS9_1A	GCF_003061925.1	19	19	7	25	0.68
DS13_1A	GCF_003061965.1	19	19	7	25	0.68
DS5_1A	GCF_003061985.1	19	19	7	25	0.68
DS11_1A	GCF_003062025.1	19	19	7	25	0.68
UBLA-34	GCF_003641085.1	19	19	7	25	0.68
s-13	GCF_013867605.1	19	19	7	25	0.68
BCRC 12255	GCF_017009485.1	19	19	7	25	0.68
BCRC 14065	GCF_017009515.1	19	19	7	25	0.68
BCRC 80064	GCF_017009695.1	19	19	7	25	0.68
LA-G80	GCF_018252545.1	19	19	7	25	0.68
LA-5	GCF_024665555.1	19	19	7	25	0.68
NBRC 13951	GCF_001591845.1	19	19	7	24	0.68
DS13_1B	GCF_003061905.1	19	19	7	24	0.68
LA_AVK1	GCF_009742735.1	19	19	7	24	0.68
s-4	GCF_013867555.1	19	19	7	24	0.68
NBIMCC 8242 (180)	GCF_021229035.1	19	19	7	24	0.68
MG-HGUT-0 2379	GCF_902386525.1	18	18	7	27	0.67
5460	GCA_021432145.1	18	18	7	26	0.67
HN017	GCF_024397395.1	18	18	7	26	0.67

LA-2	GCF_024665075.1	18	18	7	26	0.67
NC55	GCA_020883435.1	18	18	7	25	0.67
BCRC 17008	GCF_017009595.1	18	18	7	25	0.67
BCRC 17486	GCF_017009585.1	19	19	7	25	0.66
BCRC 14079	GCF_017009475.1	18	18	7	25	0.64
PNW3	GCA_004348805.1	19	19	7	23	0.63

Table 4 shows different strains of *L. acidophilus* and their ability to be a probiotic or not, results presented in descending order and generated through iProbiotics using model one, based on the criteria used by this web server tool.

Table 4. Different strains of *L. acidophilus* and their respective ability to be a probiotic according to model one of the iProbiotics web server.

Access code	Probability of being probiotic
GCF_017009485.1	99.82
GCA_000159715.1	99.80
GCF_017009475.1	99.79
GCF_017009695.1	99.79
GCF_021229035.1	99.76
GCA_017695935.1	99.75
GCF_013867605.1	99.74
GCA_017009655.1	99.74
GCA_017009575.1	99.73

GCF_017009515.1	99.73
GCA_000469765.1	99.73
GCF_003047065.1	99.72
GCA_000011985.1	99.71
GCA_000934625.1	99.70
GCA_013342945.1	99.70
GCF_000389675.2	99.70
GCF_024665555.1	99.70
GCA_000469745.1	99.69
GCF_000469705.1	99.69
GCF_002224305.1	99.69
GCA_002286215.1	99.69
GCA_001433895.1	99.67
GCF_003061905.1	99.67
GCA_000786395.1	99.67
GCF_003061985.1	99.65
GCA_004348805.1	99.64
GCF_000442865.1	99.64
GCF_017009585.1	99.63
GCF_002406675.1	99.61
GCF_017009725.1	99.60
GCF_009742735.1	99.60
GCA_000442825.1	99.59

GCA_003062045.1	99.57
GCA_003062005.1	99.57
GCF_002914945.1	99.56
GCF_008868625.1	99.54
GCF_001868765.1	99.53
GCF_003061925.1	99.53
GCA_021432145.1	99.50
GCF_024665075.1	99.50
GCA_003061945.1	99.49
GCA_020883435.1	99.49
GCA_023093425.1	99.47
GCF_001591845.1	99.47
GCA_003952845.1	99.47
GCF_001639165.1	99.46
GCF_024397395.1	99.45
GCF_017009595.1	99.44
GCF_902386525.1	99.42
GCA_017009605.1	99.30
GCA_018367455.1	99.25
GCF_018252545.1	99.24
GCF_000497795.1	99.07
GCA_003053245.1	99.02
GCF_009741835.1	99.01

GCA_017009715.1	99.00
GCA_003053135.1	98.94
GCF_003062025.1	98.89
GCA_022509485.1	98.86
GCF_013867555.1	98.86
GCF_003641085.1	98.83
GCA_003061885.1	98.51
GCF_003061965.1	97.49
GCA_900452495.1	84.26

Table 5 different strains of *L. acidophilus* and their ability to be a probiotic *Lactobacillus* or not, results presented in descending order and generated through iProbiotics using model three, based on the criteria used by this web server tool.

Table 5. Different strains of *L. acidophilus* and their respective capacity to be a probiotic according to model three of the iProbiotics web server.

Access code	Probability of being probiotic <i>Lactobacillus</i>
GCA_900452495.1	84.80%
GCA_017009715.1	51.90%
GCA_003061885.1	34.60%
GCA_018367455.1	32.70%
GCF_002914945.1	32.60%
GCF_003061985.1	31.20%
GCA_003053245.1	30.30%

GCF_001868765.1	30.00%
GCF_000497795.1	28.60%
GCF_003641085.1	27.40%
GCF_017009585.1	26.30%
GCF_018252545.1	24.90%
GCA_023093425.1	24.60%
GCF_002406675.1	23.20%
GCF_017009725.1	22.50%
GCA_017695935.1	22.10%
GCF_003061965.1	21.80%
GCF_017009595.1	21.60%
GCF_003061925.1	21.20%
GCF_013867605.1	21.00%
GCA_004348805.1	20.30%
GCF_902386525.1	20.00%
GCF_009742735.1	19.60%
GCF_017009485.1	19.60%
GCF_017009475.1	18.80%
GCA_001433895.1	18.30%
GCF_003062025.1	17.90%
GCF_017009695.1	17.70%
GCA_022509485.1	17.20%
GCF_013867555.1	17.20%

GCF_001639165.1	17.10%
GCA_017009575.1	17.00%
GCA_003061945.1	16.20%
GCA_017009655.1	14.30%
GCF_009741835.1	14.30%
GCA_003062045.1	14.00%
GCA_017009605.1	12.70%
GCF_008868625.1	12.10%
GCF_017009515.1	10.70%
GCF_021229035.1	10.70%
GCA_003053135.1	9.40%
GCF_001591845.1	9.30%
GCF_003061905.1	8.70%
GCF_024397395.1	8.60%
GCA_021432145.1	8.30%
GCF_024665075.1	8.30%
GCA_003952845.1	7.90%
GCA_020883435.1	7.20%
GCA_003062005.1	7.10%
GCA_000469745.1	5.40%
GCA_000786395.1	5.40%
GCF_000469705.1	5.30%
GCA_000469765.1	5.00%

GCA_000442825.1	4.70%
GCF_000442865.1	4.20%
GCF_002224305.1	4.00%
GCA_000011985.1	3.70%
GCA_000934625.1	3.70%
GCA_002286215.1	3.70%
GCA_013342945.1	3.70%
GCF_000389675.2	3.70%
GCF_024665555.1	3.70%
GCF_003047065.1	3.30%
GCA_000159715.1	0.16%

ANI similarity values are presentes for each pair of strains in Supplementary Data 1.

In Table 6, it is possible to observe a ranking generated from the use of the iProbiotics tool model three, BlastKOALA. Based on this ranking, we were able to list the best strains to be tested in future *in vitro* studies.

Table 6. Ranking of the ten best strains to be used for future *in vitro* studies, generated from the iProbiotics tool (model three), BlastKOALA, listing in descending order the strains with the greatest probiotic potential to be used in future studies.

Ranking position	Strain
1°	NCTC 13721
2°	DSM 20079
3°	La-5
4°	APC2845

Figure 2. Phylogenetic tree made from the 64 strains of *L. acidophilus*, with GCA_001433895.1 being the DSM 20079 strain, GCA_900452495.1 corresponding to the NCTC 13721 strain, and finally, GCA_017009715 La-5. In this analysis, these strains were evolutionarily distant.

Discussion

Among the most promising activities of probiotics, the antimicrobial stands out, helping to compete with opportunistic pathogens, inhibiting their adhesion to the mucosa. This activity is conferred through metabolic by-products (Peng et al. 2018). The formation of these by-products is dependent on population density, probiotic strains, substrate and particularities of the species (Delgado et al. 2007, Aditya et al. 2020). The genus *Lactobacillus* spp., has heterogeneous characteristics and includes species with diverse physiological and biochemical features. At present, definition of *L. acidophilus* is shown in DNA-DNA hybridization, with its CG content of species ranging from 32% to 50%, exceedingly higher than is reported for well-defined bacterial genera (Schleifer and Ludwig 1995, Bull et al. 2013).

From the 64 strains tested, *L. acidophilus* NCTC 13721, showed promising results. This strain is still poorly characterized, it was obtained from the vaginal microbiota of a volunteer patient in the United Kingdom and is available from the National Collection of Type Cultures bank (NCTC: 13721). It is already known that the genus *Lactobacillus* predominates in approximately 70% of the vaginal microbiota (Qian et al. 2021). Previous studies show that in addition to the acidophilus species, other species such as *L. iners*, *L. jensenii*, *L. crispatus* and *L. gasseri* are also present in the healthy vaginal microbiota of women, providing protection against other pathogens (Amabebe and Anumba 2018). NCTC 13721 had 90 genes related to probiotic activity, followed by the DSM 20079 strain with 77 genes. In addition, NCTC 13721 showed a significant quorum sensing (QS) compared to the other strains, as it obtained a greater number of genes involved in its metabolism. These results are important since lactic acid bacteria with probiotic capacity, the group to which *L. acidophilus* belongs, are related to the production of organic acids, bacteriocins, hydrogen peroxide and biosurfactants (Gao et al. 2022). In this way, they can act by inhibiting bacterial QS, fighting other microorganisms that are present in the same environment (Savijoki et al. 2022). The QS is fundamental in the formation of biofilm, which is capable of increasing the time that the bacteria will reside in the intestine. In this way, making the exchange of nutrients between the host and the

microbiota. Furthermore, QS is able to bring about a cooperative change in the expression of bacterial genes, such as the expression of virulence factors (Deng et al. 2020).

Regarding the genes involved in the production of lactic acid, strains NCTC13721 and DSM 20079 showed higher expression in relation to more strains. In contrast, the difference in expression ranged around one percent. Among the metabolites produced by *Lactobacillus* spp., lactic acid is an elementary antimicrobial factor (Tachedjian et al. 2017), responsible for reducing the pH, which leads to inhibition of the growth of pathogenic bacteria (Hernández-Aquino et al. 2019), since the pH of several pathogenic bacteria is slightly alkaline (Gao et al. 2022). The inhibitory and biocidal effects of pure lactic acid *in vitro*, it is able to act against Gram-negative bacteria: *Salmonella enteritidis*, *Escherichia coli* and *Pseudomonas aeruginosa*, and also against Gram-positive: *Enterococcus faecalis*, *Staphylococcus aureus*, *Bacillus cereus* and *Listeria monocytogenes* (Stanojevic et al. 2016). This inhibitory effect occurs, since lactic acid is capable of altering the pH, preventing the activity of the pathogen's urease, making this microorganism unable to grow at the adhesion site, thus acting as a bactericidal agent. This important acid is responsible for suppressing pro-inflammatory responses mediated by immune cells (Watanabe et al. 2009). It also causes intestinal and immunological homeostasis, through enterocyte renewal and macrophage mobilization (Markowiak-Kopeć et al. 2020).

In pangenome analyses, we were able to observe that as new genomes are added, the pangenome tends to increase. Thus, in this work, we evidenced an open pangenome and a stabilized core genome. Furthermore, it is visible that at some points there is discontinuity, indicating diversity in the genomes of the strains used. Confronting the data obtained in this work, in another research, in which the OrthoFinder tool was used, seven strains of *L. acidophilus* were analyzed, with C5 isolated from canine feces and the others selected because they had complete genomic information available at the NCBI (NCFM; La14 ; FSI4; ATCC53544; LA1; DSM2007). As a result, a pangenome of 2254 genes was obtained, a core genome with 1726 genes and 200 accessory genes (Son et al. 2021). Therefore, we suggest that the difference our study research and the study by Son et al. (2021), is succeeded by the significant difference in the number of genomes in each research. It is important to mention that the difference in results may occur due to the use of different tools and algorithms.

From the analyzes carried out with the online Bioinformatics tool, responsible for probiotic prediction - iProbiotics -, we noticed that it was little used until the moment of writing this

article. iProbiotics uses machine learning, not mapping probiotic action mechanisms, which are fundamental to analyze whether or not an organism has the capacity to perform such a function (Sun et al. 2021). It is already known that each probiotic will have specific characteristics, and its beneficial activities to the host, may not be the same, since each strain has its individual particularities (Rajab et al. 2020). Therefore, in this study we made a prediction probiotics based on the criteria used as probiotic mechanisms dictated by the iProbiotics program. We were able to notice that, when we compare model one with model three, the first shows approximately 99% of the strains with probiotic capacity above 90%. Therefore, it is understood that model three is more specific for *Lactobacillus* spp. with probiotic capacity, informing different percentages for each strain.

According to ANI-based molecular classification standards, values below 96% similarity and 90% global alignment indicate that the isolates may belong to different species (Ciufo et al. 2018). The NCTC13721 strain showed approximately 81% of genomic similarity when compared to the others, including La-5. Therefore, further characterization studies of the NCTC13721 strain are needed to confirm this characteristic. We highlight that, from the phylogenetic tree, the most promising strains: NCTC 13721, DSM 20079 and La-5, do not have proximity in their evolutionary characteristics, reinforcing the idea of heterogeneity. *L. acidophilus* La-5, showed about 100% genomic similarity with eight strains, namely: *L. acidophilus* DS20_1, *L. acidophilus* LA-5, *L. acidophilus* FSI4, *L. acidophilus* APC2845, *L. acidophilus* LA-G80 -111, *L. acidophilus* DS1_1A, *L. acidophilus* LA1 and *L. acidophilus* BCRC 14065. However, these strains have few reports in the literature, consequently few in vitro evaluations. *L. acidophilus* LA-5, stands out for its ability to compensate for the enteric microbiota, improves lactose digestion, and stimulates the immune response (Lin; Chen, 2000; Meng et al., 2021). Fermented milk containing this strain is able to promote protection by increasing the production of beneficial bacteria and reducing those potentially pathogenic (Savard et al., 2011; Meng et al., 2021). Thus, further studies on strains similar to La-5 are needed, as it was identified as a promising strain for future tests. However, the rest of the strains similar to it have not been reported since the writing of this article.

Therefore, we understand that the NCTC13721 and La-5 strains are the most promising for prospecting new probiotic *Lactobacillus*, and it is also pertinent to evaluate other strains genetically close to the latter, such as DS20_1, LA-5, FSI4, APC2845, LA-G80- 111, DS1_1A, LA1 and BCRC 14065. Finally, this *in silico* methodology can be improved and applied in the future to identify other probiotic organisms.

Conclusion

Based on analyzes of 64 strains of *Lactobacillus acidophilus*, using the tools of Bioinformatics Roary, ANI, BlastKoala and iProbiotics, it was possible to generate a possible ranking of the most promising strains of *Lactobacillus acidophilus* to be evaluated experimentally in future work. In addition, it was possible to deepen the knowledge about genomic similarity, pangenome, generated metabolites and ability to be a probiotic *Lactobacillus* of these different strains. With regard to the ability to be a probiotic *Lactobacillus* or not, it was not possible to perform a comparison with the iProbiotics tool, due to the data of the strains used by this program not being available. It is also important to mention that the iProbiotics results may have shown variability due to some genomes used in this study being draft, which could result in fragmented genes, making the tool less accurate.

Agradecimentos

The authors are grateful to OmixLab - Bioinformatics Laboratory of the Federal University of Pelotas. We also thank CAPES and the Graduate Program in Health Sciences for the scholarship granted, which made this work possible.

Interest conflicts

None declared.

References

- Aditya A, Peng M, Young A, Biswas D. Antagonistic Mechanism of Metabolites Produced by *Lactobacillus casei* on Lysis of Enterohemorrhagic *Escherichia coli*. *Front Microbiol.* 2020 Nov 23;11:574422. doi: 10.3389/fmicb.2020.574422.
- Agarwal G, Gitaitis RD, Dutta B. Pan-Genome of Novel *Pantoea stewartii* subsp. *indologenes* Reveals Genes Involved in Onion Pathogenicity and Evidence of Lateral Gene Transfer. *Microorganisms.* 2021; 9(8):1761. <https://doi.org/10.3390/microorganisms9081761>.

- Amabebe E, Anumba DOC. The Vaginal Microenvironment: The Physiologic Role of *Lactobacilli*. *Front Med (Lausanne)*. 2018 Jun 13;5:181. doi: 10.3389/fmed.2018.00181.
- Bull M, Plummer S, Marchesi J, Mahenthiralingam E. The life history of *Lactobacillus acidophilus* as a probiotic: a tale of revisionary taxonomy, misidentification and commercial success. *FEMS Microbiol Lett*. 2013 Dec;349(2):77-87. doi: 10.1111/1574-6968.12293.
- Carpi FM, Coman MM, Silvi S et al. Comprehensive pan-genome analysis of *Lactiplantibacillus plantarum* complete genomes. *J Appl Microbiol*. 2022 Jan;132(1):592-604. doi: 10.1111/jam.15199.
- Ciufo S, Kannan S, Sharma S et al .Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *Int J Syst Evol Microbiol*. 2018 Jul;68(7):2386-2392. doi: 10.1099/ijsem.0.002809.
- Delgado A, Arroyo LFN, Brito D et al. Optimum bacteriocin production by *Lactobacillus plantarum* 17.2b requires absence of NaCl and apparently follows a mixed metabolite kinetics. *J. Biotechnol*. 130, 193-201. Doi:10.1016/j.jbiotec.2007.01.041
- Deng Z, Luo XM, Liu J et al. Quorum Sensing, Biofilm, and Intestinal Mucosal Barrier: Involvement the Role of Probiotic. *Front Cell Infect Microbiol*. 2020 Sep 25;10:538077. doi: 10.3389/fcimb.2020.538077.
- FAO/WHO. World Health Organization. Joint FAO/WHO Consultation on Evaluation of Health and Nutritional Properties of Probiotics in Food Including Powder Milk with Live Lactic Acid Bacteria, 2001.
- Fuentes G S. A novel, sequencing-free strategy for the functional characterization of *Taenia solium* proteomic fingerprint. *PLoS Negl Trop Dis*. 2021 Feb 18;15(2):e0009104. doi: 10.1371/journal.pntd.0009104.
- Gao H, Li X, Chen X, et al. The Functional Roles of *Lactobacillus acidophilus* in Different Physiological and Pathological Processes. *J Microbiol Biotechnol*. 2022 Oct 28;32(10):1226-1233. doi: 10.4014/jmb.2205.05041.
- Goldstein J C E, Tyrrell K L. Citron D M. *Lactobacillus* Species:Taxonomic Complexity and Controversial Susceptibilities, *Clinical Infectious Diseases*, Volume 60, Issue suppl_2, May 2015, Pages S98 S107, <https://doi.org/10.1093/cid/civ072>.

Huang Z, Zhou X, Stanton C et al. Comparative Genomics and Specific Functional Characteristics Analysis of *Lactobacillus acidophilus*. *Microorganisms*. 2021 Sep 20;9(9):1992. doi:10.3390/microorganisms9091992.

Hernández-Aquino S, Miranda- Romero L A, Fujikawa H et al. Antibacterial Activity of Lactic Acid Bacteria to Improve Shelf Life of Raw Meat. *Biocontrol Sci*. 2019;24(4):185-192. doi: 10.4265/bio.24.185.

Jain C, Rodriguez-R L M, Phillippy AM. et al. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9, 5114 (2018). <https://doi.org/10.1038/s41467-018-07641-9>

Li W, Zhang Y, Li H et al. Effect of soybean oligopeptide on the growth and metabolism of *Lactobacillus acidophilus* JCM 1132. *RSC Adv*. 2020 Apr 29;10(28):16737-16748. doi: 10.1039/d0ra01632b.

Liang X, Zhu W, Lv Z et al. Molecular Computing and Bioinformatics. *Molecules*. 26;24(13):2358.2019. doi: 10.3390/molecules24132358.

Markowiak-Kopec P, Śliśewska K. The Effect of Probiotics on the Production of Short-Chain Fatty Acids by Human Intestinal Microbiome. *Nutrients*, v. 12, n. 4, p. 1-23, 16 abr. 2020. MDPI AG. <http://dx.doi.org/10.3390/nu12041107>.

Ojha A K, Shah N P, Mishra V. Conjugal Transfer of Antibiotic Resistances in *Lactobacillus* spp. *Current Microbiology*, 78(8), 2021 2839–2849. doi:10.1007/s00284-021-025541.

Peng M, Tabashsum Z, Patel P et al. Linoleic acids overproducing *Lactobacillus casei*/ limits growth, survival and virulence of *Salmonella typhimurium* and enterohaemorrhagic *Escherichia coli*. *Front. Microbiol*. 2018, 9:2663 doi: 10.3389/fmicb.2018.02663

Qian Z, Zhu H, Zhao D et al. Probiotic *Lactobacillus* sp. Strains Inhibit Growth, Adhesion, Biofilm Formation, and Gene Expression of Bacterial Vaginosis-Inducing *Gardnerella vaginalis*. *Microorganisms*. 2021 Mar 31;9(4):728. doi: 10.3390/microorganisms9040728.

Rajab S, Tabandeh F, Shahraky MK, et al. The effect of lactobacillus cell size on its probiotic characteristics. *Anaerobe*. 2020 Apr;62:102103. doi: 10.1016/j.anaerobe.2019.102103.

Saarela, M, Mogensen G, Fondén, R et al. Probiotic bacteria: safety, functional and technological properties. *J. Biotechnol*. 84, 197-215. doi: 10.1016/s0168-1656(00)00375-8.

- Savijoki K, San-Martin-Galindo P, Pitkänen K et al. Food-Grade Bacteria Combat Pathogens by Blocking AHL-Mediated Quorum Sensing and Biofilm Formation. *Foods*. 2022 Dec 24;12(1):90. doi: 10.3390/foods12010090.
- Schleifer K-H, Ludwig W (1995) Phylogeny of the genus *Lactobacillus* and related genera. *Syst Appl Microbiol*18: 461–467.
- Sitto F, Battistuzzi FU. Estimating Pangenomes with Roary. *Mol Biol Evol*. 2020 Mar 1;37(3):933-939. doi: 10.1093/molbev/msz284.
- Son S, Lee R, Park SM et al. Complete genome sequencing and comparative genomic analysis of *Lactobacillus acidophilus* C5 as a potential canine probiotics. *J Anim Sci Technol*. 2021 Nov;63(6):1411-1422. doi: 10.5187/jast.2021.e126.
- Stanojević-Nikolić S, Dimić G, Mojović L, et al. Atividade antimicrobiana do ácido láctico contra microrganismos patogênicos e deteriorantes. *J Food Process Preserv*, 2016 40: 990–998.
- Sun Y, Li H, Zheng L et al. iProbiotics: a machine learning platform for rapid identification of probiotic properties from whole-genome primary sequences. *Brief Bioinform*. 2022 Jan 813 17;23(1):bbab477. doi: 10.1093/bib/bbab477.
- Tachedjian G, Aldunate M, Bradshaw CS, et al. The role of lactic acid production by probiotic *Lactobacillus* species in vaginal health. *Res Microbiol*. 2017 Nov-Dec;168(9-10):782-792. doi: 10.1016/j.resmic.2017.04.001.
- Xiao Y, Zhao J, Zhang H et al. Mining genome traits that determine the different gut colonization potential of *Lactobacillus* and *Bifidobacterium* species. *Microb Genom*. 2021 Jun;7(6):000581. doi: 10.1099/mgen.0.000581.
- Zheng J, Wittouck S, Salvetti E et al. A Taxonomic Note on the Genus *Lactobacillus*: Description of Novel Genera, Emended Description of the Genus *Lactobacillus* Beijerinck 1901, and Union of *Lactobacillaceae* and *Leuconostocaceae*. *Int. J. Syst. Evol. Microbiol*. 2020 70 (4), 879 2782–2858. doi:10.1099/ijsem.0.004107.
- Watanabe T, Nishio H, Tanigawa T, et al. Probiótico *Lactobacillus casei* cepa *Shirota* previne lesão do intestino delgado induzida por indometacina: envolvimento do ácido láctico. *Am J Physiol Gastrointestinal Fígado Physiol*. 2009 297: G506–G513 <https://doi.org/10.1152/ajpgi.90553.2008> .

5 CONSIDERAÇÕES FINAIS

As cepas *L. acidophilus* NCTC 13721 e *L. acidophilus* DSM 20079 evidenciaram maior quantidade de genes relacionados a atividade metabólica, apresentando *scores* maiores se comparados as outras cepas. Sendo assim, futuramente serão realizados testes *in vitro* para corroborar com os achados deste trabalho.

São necessárias pesquisas mais aprofundadas sobre as cepas: DS20_1, LA-5, FSI4, APC2845, LA-G80-111, DS1_1A, LA1 e BCRC 14065, visto que possuem similaridade genômica com *L. acidophilus* La-5, que mostrou resultados promissores.

Os achados relacionados a análise pangenômica, evidenciaram um pangenoma aberto e um genoma núcleo estabilizado. Conseguimos notar também que em alguns pontos há descontinuidade, indicando uma diversidade nos genomas das cepas utilizadas, fator que pode ocorrer devido ao gênero *Lactobacillus* spp. ser heterogêneo.

Por fim, os resultados gerados neste estudo poderão ser utilizados para propor futuros testes *in vitro* com as cepas mais promissoras, a fim de avaliar sua inibição contra microrganismos patogênicos e seus benefícios ao hospedeiro.