



Universidade Federal do Pará

Instituto de Ciências Biológicas

Programa de Pós-Graduação em Genética e Biologia Molecular

ELIANE EVANOVICH BARBOSA DOS SANTOS

**GENÔMICA COMPARATIVA DE DUAS BACTÉRIAS DA MICROBIOTA
INTESTINAL *BACTEROIDES FRAGILIS* E *LACTIPLANTIBACILLUS*
*PLANTARUM***

Belém- PA

Dezembro/2021



Universidade Federal do Pará

Instituto de Ciências Biológicas

Programa de Pós-Graduação em Genética e Biologia Molecular

**GENÔMICA COMPARATIVA DE DUAS BACTÉRIAS DA MICROBIOTA
INTESTINAL *BACTEROIDES FRAGILIS* E *LACTIPLANTIBACILLUS*
*PLANTARUM***

ELIANE EVANOVICH BARBOSA DOS SANTOS

Tese submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFPA como requisito parcial para obtenção do grau de Doutor em Genética e Biologia Molecular.

Orientador: Dr. João Farias Guerreiro (UFPA)

Belém- PA

Dezembro/2021

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)

B238g Barbosa Evanovich dos Santos, Eliane.
 Genômica Comparativa de duas bactérias da microbiota
 intestinal *Bacteroides fragilis* e *Lactiplantibacillus plantarum* /
 Eliane Barbosa Evanovich dos Santos. — 2021.
 132 f. : il. color.

 Orientador(a): Prof. Dr. João Farias Guerreiro
 Tese (Doutorado) - Universidade Federal do Pará, Instituto de
 Ciências Biológicas, Programa de Pós-Graduação em Genética e
 Biologia Molecular, Belém, 2021.

 1. *Bacteroides fragilis*. 2. *Lactiplantibacillus plantarum*. 3.
 Genômica comparativa. 4. Resistência a antibióticos. 5.
 Fatores de virulência. I. Título.

CDD 579.135

ELIANE EVANOVICH BARBOSA DOS SANTOS

**GENÔMICA COMPARATIVA DE DUAS BACTÉRIAS DA MICROBIOTA
INTESTINAL *BACTEROIDES FRAGILIS* E *LACTIPLANTIBACILLUS*
*PLANTARUM***

**Tese aprovada como requisito para obtenção do grau de Doutor no
curso de Programa de Pós-Graduação em Genética e Biologia
Molecular da UFPA pela comissão formada pelos seguintes membros:**

Orientador:

Prof. Dr. João Farias Guerreiro

Universidade Federal do Pará (Titular)

Profª. Dra. Renata Coelho Rodrigues Noronha

Universidade Federal do Pará (Titular)

Prof. Dr. Bruno Rafael Ribeiro de Almeida

Instituto Federal de Educação, Ciência e Tecnologia do Pará (Titular)

Prof. Dr. Rommel Thiago Jucá Ramos

Universidade Federal do Pará (Titular)

Prof. Dr. Luis Carlos Guimarães

Universidade de Brasília (Titular)

Profª. Dra. Joana Montezano Marques

Universidade Federal do Pará (Suplente)

À memória de meus queridos pais Manoel e
Maria Joaquina e de meu amado marido
Mário Emanuel.

AGRADECIMENTOS

Agradeço primeiramente aos meus irmãos Bete, Marcos, Cláudio e Lúcia; cunhados Moisés e Stella; sobrinhos Ian, Pilar, Chiara, Ivan, Andrei e Clarinha; e primos, Bena, Ednilson, Júlio e Rosa. A todos agradeço o apoio recebido durante a realização desse trabalho. Às tias Lena e Salomé; e enteados Vitor e Carina que mesmo distantes, me apoiaram incondicionalmente.

Agradeço aos amigos Jackson, Sheila, Nilson, Gleomar, Renata, Mônica, Leandro, Verônica, Amanda e Wilcley por toda a compreensão no ambiente de trabalho, nos piores e melhores momentos.

Às amigas Patrícia e Selma pelo apoio e troca de ideias e carinho.

Aos professores da Pós-graduação em Genética que compartilharam o seu conhecimento conosco e com isso tornaram possível a realização deste trabalho.

Ao meu orientador, D. João Farias Guerreiro por toda a compreensão e confiança que depositou em mim.

À Pró-Reitoria de Pesquisa e Pós-Graduação, principalmente à Prof.^a Dra. Maria Iracilda da Cunha Sampaio, pelo apoio durante as publicações.

À equipe da secretaria da Pós-graduação por todas as dúvidas esclarecidas.

SUMÁRIO

1	INTRODUÇÃO	12
1.1	BACTÉRIAS GRAM-POSITIVAS E GRAM-NEGATIVAS	12
1.2	A MICROBIOTA INTESTINAL HUMANA	14
1.3	A MICROBIOTA INTRAUTERINA	16
1.4	A MICROBIOTA INTESTINAL E IMUNIDADE	17
1.5	<i>BACTEROIDES FRAGILIS</i>	18
1.5.1	Fragilisina	19
1.5.2	Perforina	22
1.5.3	Don PUL	23
1.5.4	<i>Locus</i> de Biossíntese de Polissacarídeo Capsular	24
1.5.5	Sistema Secretor Tipo VI (T6SS)	24
1.5.6	Genes de Resistência a Antibióticos	26
1.6	LACTOBACILLACEAE	28
1.6.1	LAB e Probióticos	30
1.6.2	<i>Lactiplantibacillus plantarum</i>	31
1.6.3	Propriedades Genômicas de <i>L. plantarum</i>	32
1.6.4	Bacteriocinas de <i>L. plantarum</i>	33
1.7	GENÔMICA COMPARATIVA	35
1.7.1	Alinhamento Genômico	36
1.7.2	Pan-genoma	37

1.7.3	Sequências CRISPR-Cas	39
1.7.4	Predição do Mobiloma	41
1.7.5	Genes de Resistência a Antibióticos	43
1.7.6	Genes de Virulência	45
1.7.7	Análise de Evolução Molecular	46
1.8	OBJETIVOS	49
1.8.1	Objetivo Geral	49
1.8.2	Objetivos Específicos	49
2.	CAPÍTULO I: COMPARATIVE GENOMICS ANALYSIS OF <i>BACTEROIDES FRAGILIS</i>	50
3.	CAPÍTULO II: COMPARATIVE GENOMIC ANALYSIS OF <i>LACTOBACILLUS PLANTARUM</i>: AN OVERVIEW	76
4	DISCUSSÃO	88
	<i>Bacteroides fragilis</i>	88
	<i>Lactiplantibacillus plantarum</i>	89
	Contrastes Genômicos	90
	Programas utilizados nas análises genômicas	91
	Impacto Acadêmico das Publicações	92
5	CONSIDERAÇÕES FINAIS	96
	BIBLIOGRAFIA	97
	APÊNDICE	124

LISTA DE ABREVIATURAS

BFT- Frangilisina

CARD - Banco de dados abrangente de resistência a antibióticos

COG- Clusters of Orthologous Groups of proteins

Ctn – Transposons Conjugativos

ETBF - *Bacteroides fragilis* enterotoxigênicas

HGT- Transferência horizontal de genes

ICE - Elementos Integrativos e Conjugativos

IME - Elementos integrativos e mobilizáveis

IS - Elementos de inserção

KEGG - Kyoto Encyclopedia of Genes and Genomes

LAB - Bactérias do ácido láctico

LPS - Locus de biossíntese de polissacarídeo

MGE - Elementos genéticos móveis

NTBF - *Bacteroides fragilis* não-toxigênicas

T4SS - Sistemas de secreção bacteriana do tipo IV

T6SS - Sistema Secretor Tipo VI

Tn – Transposon

VFDB - Banco de dados de fatores de virulência

LISTA DE FIGURAS

- Figura 1:** Comparação entre os envelopes celulares de bactérias gram-positivas em A) e gram-negativas em B). Modificado de Kayser et al. (2005, pp.155-156). 13
- Figura 2:** Mudanças na diversidade da microbiota intestinal em diferentes fases da vida causadas por perturbações fisiológicas (nutrição, desnutrição, obesidade e uso de antibióticos) de acordo com dados do gene ribossomal 16S (simbolizada na figura por 16S e análise metagenômica (simbolizada por DNA). Adaptado de Ottman *et al.* (2012). 15
- Figura 3:** Alteração na estrutura da proteína BFT. O gene *bft* codifica uma holotoxina pré-pró-proteína que é processada por diferentes enzimas e apresenta como produto final uma proteína madura de 20 kDa. Fonte: Sears; Geis; Housseau (2014). 20
- Figura 4:** Mecanismo de ação da BFT nos colonócitos. A enzima se liga a um receptor específico (não caracterizado), cliva a E-caderina e desencadeia a secreção de citocinas pró-inflamatórias que contribuem para aspectos-chave do potencial carcinogênico de ETBF. A clivagem da E-caderina também libera a liberação de β -catenina que estimula a transcrição do proto-oncogene *c-myc*. Adaptado de Sears; Geis; Housseau (2014). Informação adicional sobre c-Myc retirada de Ko *et al.* (2016). 22
- Figura 5:** Mapeamento das ORF (do inglês *open reading frame*, ou Quadro de leitura aberto) no *locus GA3 T6SS* de duas linhagens de *B. fragilis* (638R e 9343). As linhas vermelhas indicam regiões com identidade igual ou maior que 95%. Os genes das regiões variáveis são mostrados em branco e as linhas verdes correspondem às regiões deletadas nos experimentos que ajudam a definir as funções gênicas. Adaptado de Chatzidaki-Livanis; Geva-Zatorsky; Comstock (2016). 26
- Figura 6:** Modelo do *locus pln* das linhagens C11, WCFS11 e V90. A organização é similar, mas i) V90 possui um transposon simbolizado pelo triângulo, e *plnQ*, *plnS* e *plnT* aparecem como um único gene na imagem; iii) em C11, o gene *plnXY* está ausente na região à jusante de *plnW*. Os círculos indicam promotores regulados. Adaptado de Diep *et al.* (2009). 35

RESUMO

Cerca de 80% das bactérias da microbiota intestinal humana de adultos saudáveis pertencem aos filos Bacteroidetes e Firmicutes. Vários estudos revelaram que a relação Firmicutes/Bacteroidetes é essencial à homeostase intestinal. Devido à grande relevância dessas bactérias, analisamos os genomas de dois representantes delas, *Bacteroides fragilis* (Bacteroidetes) e *Lactiplantibacillus plantarum* (Firmicutes) com o objetivo de fornecer mais informações sobre sua evolução genômica. Utilizamos diferentes tipos de ferramentas de genômica comparativa que mostraram que os elementos genéticos móveis, como CtnDOT, CTn341 e ISBf1, parecem ter grande importância na disseminação dos genes de resistência a antibióticos nos genomas de *B. fragilis*. Também encontramos evidências de transferência horizontal de genes entre uma região de *B. fragilis* DCMOUH0085B e *Butyrivibrio faecalis* H184. Essas regiões genômicas são ICE putativas com T4SS e carregam os genes de resistência a antibióticos, *aadE* e *ermB*. A seleção positiva foi detectada em 398 genes de *B. fragilis*, incluindo os genes *cepA* e *tetQ*. Além disso, o gene *bft* que codifica a endotoxina fragilisina, pode ser disseminado entre as linhagens de *B. fragilis* através do transposon Ctn86. No entanto, sua função parece ter sido perdida em muitas linhagens. 76,9% das linhagens de *B. fragilis* possuem o sistema CRISPR-Cas, em contraste, apenas 18% das *L. plantarum* analisadas nesta pesquisa possuem este sistema. A domesticação de elementos genéticos móveis, incluindo bacteriófagos, foi descrita em outros *Lactiplantibacillus*, e esse processo pode incluir *L. plantarum*. Nossos resultados também indicaram que as linhagens de *L. plantarum* podem produzir diferentes vitaminas do complexo B, que auxiliam no enriquecimento nutricional, além das bacteriocinas lactobina, pediocina e PLNC8 $\alpha\beta$, que somadas às plantaricinas e nisina, auxiliam na preservação de alimentos e podem garantir utilizações específicas de cada linhagem na fermentação de alimentos ou na indústria farmacêutica. Os genomas de *L. plantarum* não possuem genes de resistência a antibióticos ou determinantes de virulência, sendo uma bactéria segura para uso comercial.

Palavras-chave: *Bacteroides fragilis*, *Lactiplantibacillus plantarum*, genômica comparativa, pan-genoma, Crispr-Cas, resistência a antibióticos, fatores de virulência, evolução molecular.

ABSTRACT

Approximately 80% of the bacteria of the human gut microbiota from healthy adults belong to Bacteroidetes and Firmicutes. Several studies have revealed that the Firmicutes/Bacteroidetes ratio is essential to intestinal homeostasis. Due to the great relevance of these bacteria, we analyzed the genomes of two representatives of them, *Bacteroides fragilis* (Bacteroidetes) and *Lactiplantibacillus plantarum* (Firmicutes), to provide further information about their genomic evolution. We used different types of comparative genomics tools which indicated that mobile genetic elements, such as CtnDOT, CTn341 and ISBf1 seem to have considerable importance in the spread of genes for resistance to antibiotics in the *B. fragilis* genomes. We also detected evidence of the horizontal gene transfer between a region from the *B. fragilis* DCMOUH0085B strain and ones from *Butyricimonas faecalis* H184. These genomic regions are putative ICE with T4SS and carry the antibiotic resistance genes, *aadE* and *ermB*. Positive selection was detected in 398 genes from *B. fragilis*, including the *cepA* and *tetQ* genes. Beyond that, the *bft* gene, which encodes the endotoxin fragilysin, may be disseminated among the *B. fragilis* strains through the transposon Ctn86. Nevertheless, its function seems to have been lost in many strains. 76.9% of *B. fragilis* strains have the CRISPR-Cas system, in contrast, only 18% of *L. plantarum* analyzed in this research have this system. The domestication of mobile genetic elements, including bacteriophages, has been described in other *Lactiplantibacillus*, and this process may include *L. plantarum*. Our research also indicated that *L. plantarum* strains produce different complex B vitamins, which helps in nutritional enrichment; in addition to bacteriocins lactobin, pediocin, PLNC8 $\alpha\beta$. These bacteriocins, added to plantaricins and nisins, help in food preservation and can guarantee specific uses for each strain, in food fermentation or into the pharmaceutical industry. Furthermore, *L. plantarum* genomes do not have antibiotic resistance genes or virulence determinants, being a bacterium safe for commercial use.

Keywords: *Bacteroides fragilis*, *Lactiplantibacillus plantarum*, comparative genomic, pan-genome, Crispr-Cas, multidrug-resistant genes, molecular evolution.

1. INTRODUÇÃO

Bacteroides fragilis e *Lactiplantibacillus plantarum* são organismos com características contrastantes e pertencem aos filos mais abundantes da microbiota intestinal humana adulta, Bacteroidetes (gram-negativa) e Firmicutes (gram-positiva), respectivamente. O conceito adotado para considerar uma espécie nesta tese foi o implementado por Konstantinidis e Tiedje (2005) que caracteriza uma espécie bacteriana como um grupo de linhagens que compartilha identidade média global de nucleotídeos igual ou superior a 95%.

1.1 BACTÉRIAS GRAM-POSITIVAS E GRAM-NEGATIVAS

A coloração de gram é um dos métodos clássicos mais usados em microbiologia. Ele foi criado em 1884 por Ham Christian Gram. O método consiste em coloração inicial com violeta cristal, descoloração com solventes orgânicos (etanol e acetona) e contra-coloração, utilizando safranina (CHASKES; AUSTIN, 2015, p. 46).

A parede celular das bactérias gram-positivas e gram-negativas é formada por peptídeoglicanas (mureína). É espessa nas gram-positivas, por isso elas ficam violetas após a coloração; enquanto que a parede das bactérias gram-negativas é mais estreita e ficam rosadas após a segunda coloração (CHASKES; AUSTIN, 2015, p. 46). Uma das vantagens do método é que as bactérias podem ser analisadas por microscopia óptica.

Os dois tipos bacterianos acumulam outras diferenças, que incluem a composição da parede celular e número de membranas (KAYSER *et al.*, 2005, pp.152-156). As principais diferenças são mostradas na Figura 1.

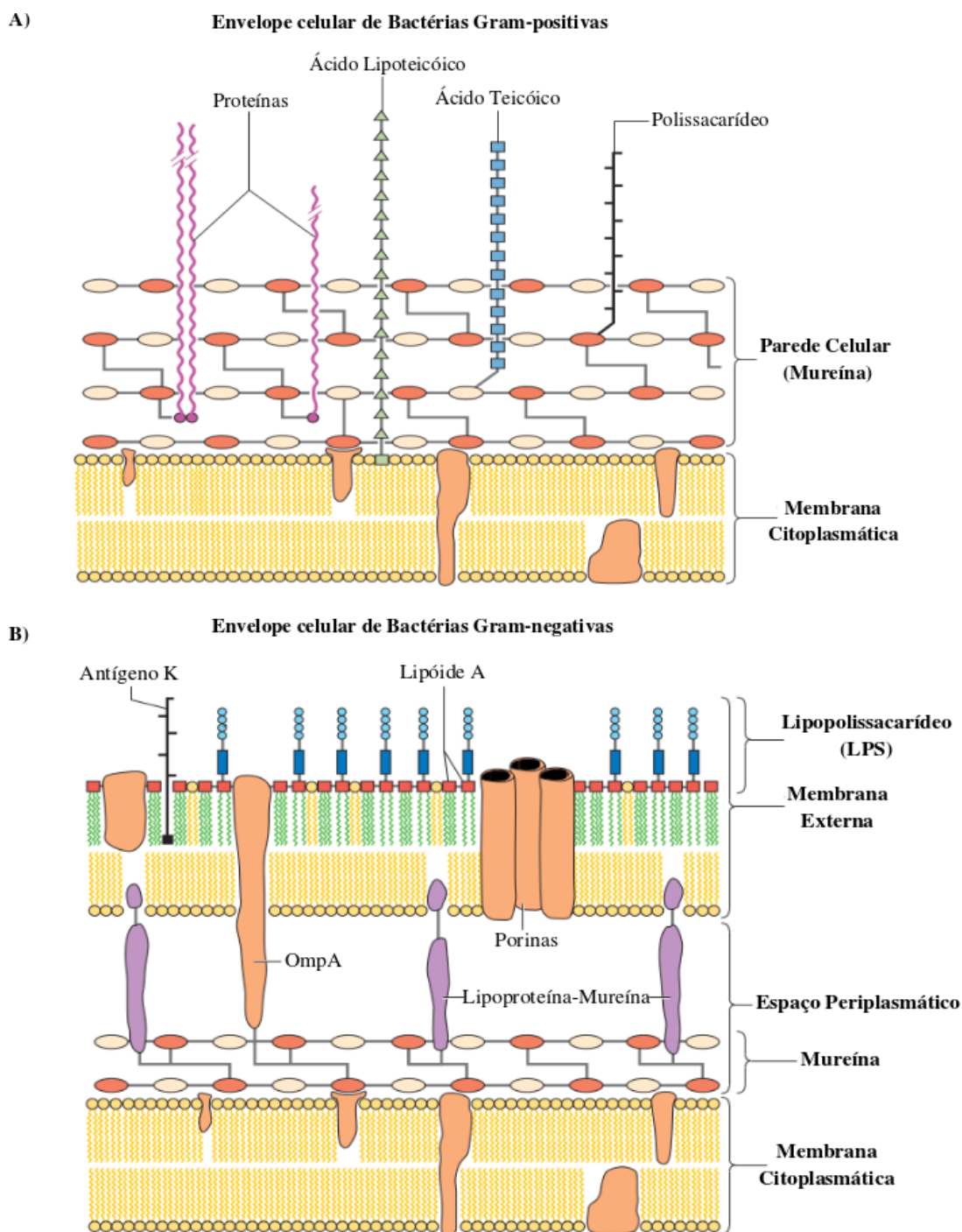


Figura 1: Comparação dos envelopes celulares de bactérias gram-positivas A) e gram-negativas B). Modificado de Kayser *et al.* (2005, pp.155-156).

1.2 A MICROBIOTA INTESTINAL HUMANA

Os intestinos delgado e grosso possuem vírus, bacteriófagos, bactérias comensais e oportunistas, fungos e protozoários, que juntos compõem a microbiota intestinal, um órgão virtual com cerca de um a dois quilogramas (SEKIROV *et al.*, 2010; PATTERSON *et al.*, 2016; CLEMENTE *et al.*, 2012). Muitas bactérias intestinais possuem função de destaque no hospedeiro, codificando proteínas essenciais à fermentação de nutrientes não digeríveis, metabolismo de drogas e toxinas, produção de nutrientes (como vitaminas do complexo B e ácidos graxos de cadeia curta -AGCC), modulação epigenética e imunidade (PATTERSON *et al.*, 2016; CLEMENTE *et al.*, 2012; JIA *et al.*, 2008; MACFARLANE; MACFARLANE, 2011; AW; FUKUDA, 2015; JEFFERY; O'TOOLE, 2013; SHEN *et al.*, 2021). A interação entre a microbiota e o hospedeiro é bastante consistente e processos vitais como a replicação e o reparo de DNA podem ser regulados por genes microbianos (AW; FUKUDA, 2015; HOOPER; MIDTVEDT; GORDON, 2002). O folato, por exemplo (vitamina B9) é produzido amplamente por bactérias do gênero *Bifidobacterium* e está envolvido nas vias metabólicas de nucleotídeos, aminoácidos, vitaminas e síntese do grupo metil (POMPEI *et al.*, 2007).

A microbiota intestinal é formada por aproximadamente 500 a 1000 espécies e mais de 7000 linhagens bacterianas (LEY; PETERSON; GORDON, 2006), na maioria dos filos Bacteroidetes e Firmicutes, com proporções reduzidas de bactérias dos filos Actinobactéria, Proteobactéria, Fusobactéria e Verrucomicrobia (ECKBURG *et al.*, 2005; XU *et al.*, 2007; LAY; DORÉ; RIGOTTIER-GOIS, 2007).

A diversidade da microbiota intestinal é estabilizada entre os dois aos três anos de idade com a predominância de bifidobactérias (KOENIG *et al.*, 2012; OTTMAN *et al.*

2012; SCHOLTENS *et al.*, 2012); atinge maior complexidade durante a fase adulta (RAJILIĆ-STOJANOVIĆ *et al.*, 2009; BIAGI *et al.*, 2010; OTTMAN *et al.* 2012; YATSUNENKO *et al.*, 2012); e menor diversidade em idosos, com aumento nas proporções de Bacteroidetes e Proteobactérias e redução do gênero *Bifidobacterium* (RAJILIĆ-STOJANOVIĆ *et al.*, 2009; BIAGI *et al.*, 2010; OTTMAN *et al.* 2012; YATSUNENKO *et al.*, 2012).

Bebês alimentados exclusivamente com leite materno possuem nos intestinos maior proporção dos gêneros *Lactobacillus* e *Bifidobacterium*, enquanto que àqueles alimentados com fórmulas infantis apresentam maior abundância de *Clostridium spp.*, *Bacteroides spp.* e membros da família Enterobacteriaceae (FALLANI *et al.* 2011; HARMSSEN *et al.*, 2007; KOROPATKIN *et al.*, 2012; MA *et al.*, 2020). As alterações na microbiota durante a vida são mostradas na Figura 2.

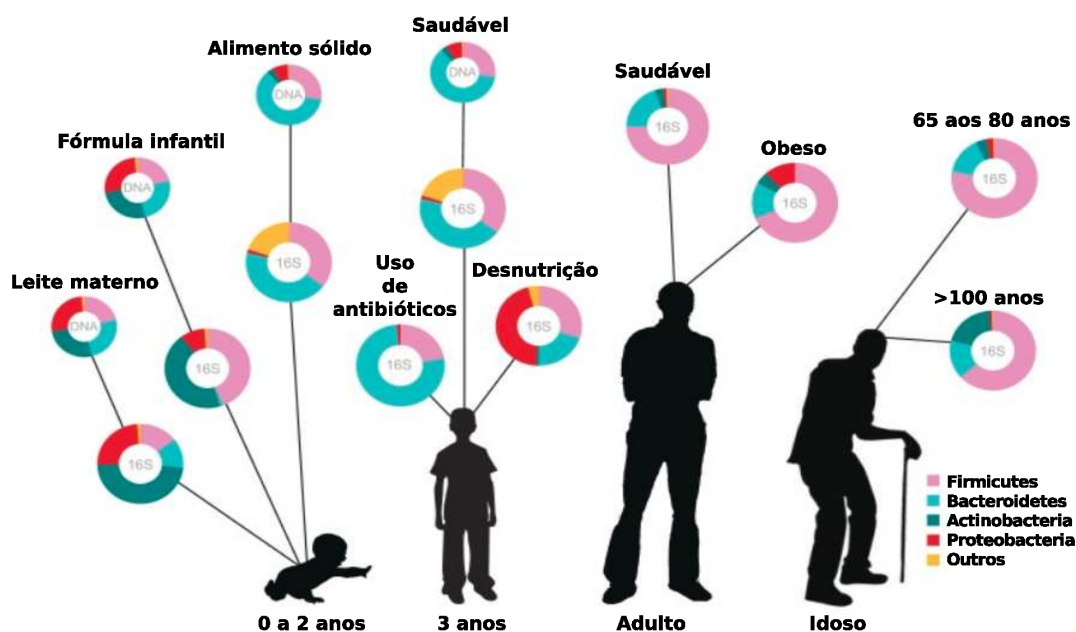


Figura 2: Mudanças na diversidade da microbiota intestinal em diferentes fases da vida causadas por perturbações fisiológicas (nutrição, desnutrição, obesidade e uso de antibióticos) de acordo com dados do gene ribossomal 16S (16S rRNA) e análise metagenômica (simbolizada por DNA). Adaptado de Ottman *et al.* (2012).

1.3 A MICROBIOTA INTRAUTERINA

O ambiente uterino não é estéril como acreditava-se até meados de 1980. Desde então, vários estudos com cultura de células indicaram a presença de bactérias no útero, mesmo em mulheres saudáveis (HEINONEN *et al.*, 1985; ESCHENBACH *et al.*, 1986; HEMSELL *et al.*, 1989; MOLLER *et al.*, 1995). Análises posteriores de placenta, líquido amniótico e mecônio também revelaram a presença de uma microbiota (PEREZ-MUÑOZ *et al.*, 2017; STINSON *et al.*, 2019). Estudos baseados em 16S mostraram que o líquido amniótico, placenta e mecônio possuem as bactérias *Lactobacillus*, *Bifidobacterium*, *Bacteroides* e *Clostridium leptum* (PEREZ-MUÑOZ *et al.*, 2017).

Além de DNA bacteriano, Stinson *et al.* (2019) identificaram citocinas inflamatórias no líquido amniótico e imunomoduladores como acetato e propionato no mecônio. Esses resultados indicam que o desenvolvimento do sistema imune tem início no ambiente intrauterino e pode ser modulado pela sua microbiota local. Além das bactérias presentes no útero, outras que formarão a microbiota intestinal do recém-nascido são também adquiridas durante o parto e variam conforme a forma de nascimento. Se o nascimento a por parto normal, a microbiota do recém-nascido terá ainda *Enterobacter*, *Enterococcus*, *Prevotella* e *Ruminococcus* (FAVIER *et al.*, 2002; PAUL *et al.*, 2015; e se por cesariana, haverá a presença de *Corynebacterium* ou outras bactérias presentes na pele da mãe e ambiente hospitalar (DOMINGUEZ-BELLO *et al.*, 2010). O leite materno parece ser um modulador da microbiota infantil, sendo metabolizado principalmente por *Bifidobacterium* (XIONG *et al.*, 2021). A composição da microbiota muda após a ingestão de alimentos sólidos, sendo influenciada por dieta, doenças adquiridas, intervenções cirúrgicas, uso de antibióticos, etc. (FAVIER *et al.*, 2002; PAUL *et al.*, 2015).

1.4 A MICROBIOTA INTESTINAL E IMUNIDADE

Segundo Gaskins (1997) a microbiota intestinal é fundamental ao desenvolvimento de componentes inatos e adquiridos do sistema imune da mucosa intestinal. Em mamíferos, esta mucosa realiza a mediação da homeostase entre a microbiota e o hospedeiro, prevenindo a ocorrência de inflamações ou sepse (GASKINS, 1997). No entanto, essa relação é bastante instável, devido a sensibilidade das células hospedeiras e a natureza antigênica da mucosa intestinal (DEPLANCKE; GASKINS, 2001). *Bacteroides fragilis* possui um polissacarídeo (PSA) que estimula a maturação e o desenvolvimento do sistema imune durante a colonização bacteriana (MAZMANIAN *et al.*, 2005). Cobaias gnotobióticas são deficientes na produção de linfócitos T CD4+, e a produção de PSA por *B. fragilis*, uma molécula imunomoduladora, é capaz de suprir essa deficiência através de diferenciação de células T reguladoras (Treg CD4+Foxp3+) que previne e cura de inflamação da mucosa intestinal (MAZMANIAN; ROUND; KASPER, 2008; ROUND; MAZMANIAN, 2010).

Durante disbioses (alteração na composição da microbiota causada pelo uso de antibióticos e imunossupressores, por estresse ou disfunção gastrointestinal) pode ocorrer a migração de bactérias ou toxinas bacterianas do lúmen intestinal para linfonodos mesentéricos e outros órgãos (translocação bacteriana) resultando em bacteremia e sepse (WIEST; RATH, 2003; MAGNOTTI; DEITCH, 2005; GATT; REDDY; MACFIE, 2007; KRENTZ; ALLEN, 2017).

1.5 *BACTEROIDES FRAGILIS*

Os filos Bacteroidetes e Firmicutes correspondem a 80% da microbiota intestinal de adultos saudáveis (OTTMAN *et al.*, 2012; RIOS-COVIAN *et al.*, 2017; VALGUARNERA; WARDENBURG, 2020). Estes filos ocupam diferentes nichos na microbiota e sua proporção parece ter impacto na saúde do hospedeiro (JOHNSON *et al.*, 2017). Bacteroidetes é o maior produtor intestinal de acetato e propionato, que podem ativar o sistema nervoso parassimpático e a lipogênese, de forma a induzir a secreção de insulina estimulada por glicose podendo originar resistência à insulina (MACFARLANE; MACFARLANE, 2003; PERRY *et al.*, 2016; LEE; LEE, 2020). Por outro lado, Firmicutes produz butirato que inibe a enzima histona desacetilase e melhora a sensibilidade à insulina em murinos (VRIEZE *et al.*, 2012; BOUTER *et al.*, 2018; LEE; LEE, 2020).

Bacteroides é um dos gêneros mais abundantes na microbiota intestinal humana e corresponde a cerca de 25% das bactérias colônicas (SALYERS, 1984; LEY; PETERSON; GORDON, 2006). É um bastonete Gram-negativo, anaeróbico que não forma esporo e é composto por 22 espécies na maioria isoladas de fezes (WEXLER, 2007). As espécies mais encontradas em amostras clínicas são *B. fragilis*, *B. distasonis*, *B. ovatus*, *B. thetaiotaomicron*, *B. vulgatus* e *B. uniformis* (WEXLER, 2007; YEKANI *et al.*, 2020).

B. fragilis corresponde a 0.1 a 0.5% das bactérias intestinais, mas é isolada em cerca de 41% das infecções intra-abdominais (TAN *et al.*, 2017; VALGUARNERA; WARDENBURG, 2020; ELSAGHIR; REDDIVARI, 2020). É comumente isolada de infecções mono e poli microbiais da cavidade abdominal, levando a óbito aproximadamente 19% dos pacientes infectados (WAREHAM *et al.*, 2005; WEXLER, 2007; GOLDSTEIN, 1996; YEKANI *et al.*, 2020). Estas infecções podem evoluir para

bacteremia e sepse; mas são condições excepcionais que dependem da vulnerabilidade do paciente, que geralmente é muito jovem (crianças de um a cinco anos) ou portador de doença crônica (doença inflamatória intestinal, cânceres, imunodeficiência, insuficiência cardíaca congestiva ou diabetes) (WAREHAM *et al.*, 2005; YOSHINO *et al.*, 2012; CHOI *et al.*, 2016; NAKAMURA *et al.*, 2017; YEKANI *et al.*, 2020; ELSAGHIR; REDDIVARI, 2020). Cerca de 4 a 20% dos portadores de enterotoxigênicas (ETBF) são assintomáticos (SAN JOAQUIN *et al.*, 1995; ZHANG *et al.*, 1999). *B. fragilis* possui vários genes e *loci* fundamentais à sua permanência e comportamento patobionte na microbiota colônica. Alguns desses genes são explanados nos tópicos abaixo.

1.5.1 Fragilisina

O primeiro relato de que algumas linhagens de *B. fragilis* podem ser toxigênicas foi descrito por Myers e colaboradores em 1984, que observaram o acúmulo de fluido em alças intestinais de cordeiros recém-nascidos com diarreia. No ano seguinte (1985), Myers e Shoop mostraram que um filtrado com peso molecular de 19,5 quilodalton (kDa) presente nas fezes dos cordeiros doentes estimulava a secreção e a ruptura de ligamentos intestinais. Posteriormente, as linhagens toxigênicas foram associadas também à diarreias agudas em crianças de um a cinco anos em diferentes países (MYERS *et al.*, 1987; SACK *et al.*, 1994; SEARS *et al.*, 2005; NGUYEN *et al.*, 2005; PATHELA *et al.*, 2005; HASAN *et al.*, 2006). A enterotoxina presente no filtrado foi chamada fragilisina (BFT) e pertence ao grupo das metaloproteases dependentes de zinco e apresenta um motivo conservado HEXXHXXGXXH, típico da subfamília Metzincina (MONCRIEF *et al.*, 1995). Essa protease cliva o domínio extracelular das E-caderinas presentes na junção aderente entre os

colonócitos do epitélio intestinal causando lesões (WU *et al.*, 1998). O gene da fragilisina (*bftP*) codifica uma holotoxina de 397 aminoácidos (aa) e apresenta três isoformas descritas BFT-1, BFT-2 e BFT-3 que diferem entre si por poucas mudanças de aminoácidos (92 a 96% de identidade) (KATO *et al.*, 2000; SEARS, 2001). A pré-pró-fragilisina é clivada entre os resíduos arginina-alanina e uma proteína madura com 186 aas é liberada (SEARS, 2001). Este processamento pós-traducional pode ocorrer por meio de proteases intestinais ou na corrente sanguínea pela ação da proteína fragipain (C11 clostripain) (SEARS, 2001; HERROU *et al.* 2016). Moncrief e colaboradores (1995) demonstraram *in vitro* que a fragilisina purificada consegue hidrolisar gelatina, actina, tropomiosina, fibrinogênio e também se auto digere. O esquema para o processamento da fragilisina é mostrado na Figura 3.

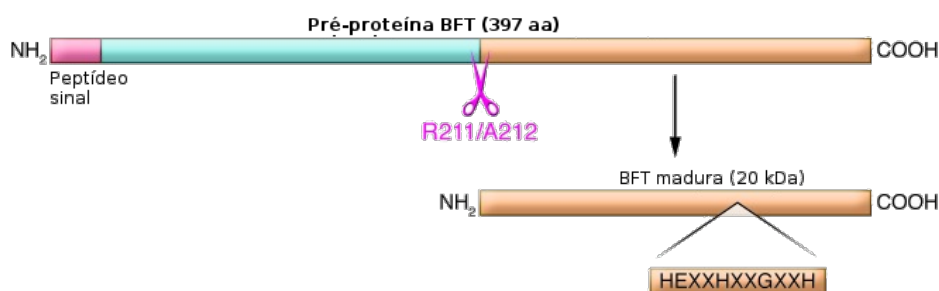


Figura 3: Alteração na estrutura da proteína BFT. O gene *bftP* codifica uma holotoxina pré-pró-proteína que é processada por diferentes enzimas e apresenta como produto final uma proteína madura de 20 kDa. Fonte: Sears; Geis; Housseau (2014).

O gene *bftP* está localizado em uma ilha de patogenicidade com cerca de 6 quilobases (kb), juntamente com os genes que codificam a metaloproteinase II e uma proteína hipotética (MONCRIEF *et al.*, 1998; FRANCO *et al.*, 2004; SHIRYAEV *et al.*, 2014).

De acordo com a produção ou não de BFT, as linhagens de *B. fragilis* são classificadas em enterotoxigênicas (ETBF) e não toxigênicas (NTBF), respectivamente.

Somente 4 a 20% dos portadores de EBTF são assintomáticos (SAN JOAQUIN *et al.*, 1995; ZHANG *et al.*, 1999).

Em 2003, Wu *et al.* descreveram a ação e as consequências da clivagem de E-caderinas. Neste experimento, eles mostraram que as β -cateninas são ligadas ao domínio intracelular das E-caderinas. Após a clivagem das E-caderinas pela fragilisina, as β -cateninas ficam livres e podem migrar para o citoplasma ou núcleo das células intestinais. No núcleo, as β -cateninas estimulam a transcrição do proto-oncogene *c-myc*. A proteína traduzida, c-Myc, é um estimulador de proliferação celular e impulsiona a produção de citocinas pró-inflamatórias (interleucina 8- IL-8) pelas células do epitélio intestinal. Essa resposta inflamatória pode aumentar a apoptose de células epiteliais facilitando a invasão bacteriana extraintestinal (KO *et al.*, 2016).

O processo inflamatório crônico estimulado pela fragilisina pode resultar em hiperplasia colônica e tumorigênese (HARDY; MELTZER; JANKOWSKI, 2000; WU *et al.*, 2003). Além disso, a liberação de BFT também sinaliza a ação de estimuladores de estresse celular (via de sinalização P38 MAPK e fator nuclear kappa B – NF- κ B), que atuam na transcrição de genes de citocinas pró-inflamatórias (SEARS; GEIS; HOUSSEAU, 2014). A Figura 4 indica os diferentes modos de ação da enzima BFT.

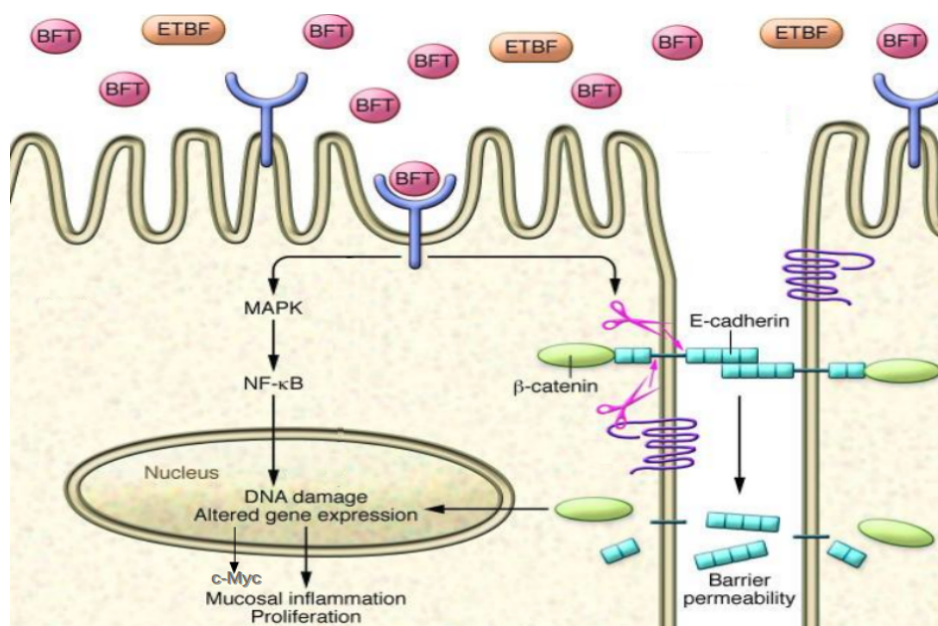


Figura 4: Mecanismo de ação da BFT nos colonócitos. A enzima se liga a um receptor específico (não caracterizado), cliva a E-caderina e desencadeia a secreção de citocinas pró-inflamatórias que contribuem para aspectos-chave do potencial carcinogênico de ETBF. A clivagem da E-caderina também libera a liberação de β -catenina que estimula a transcrição do proto-oncogene *c-myc*. Adaptado de Sears; Geis; Housseau (2014). Informação adicional sobre *c-Myc* retirada de Ko *et al.* (2016).

1.5.2 Perforina

A competição entre microrganismos da microbiota intestinal é intensa e depende de uma série de moléculas antimicrobianas, como exemplo as bacteriocinas lanticinas, lantibióticos, plantaricinas, colicinas. O mecanismo de ação entre elas é bastante diferenciado e pode envolver a inibição enzimática, liberação de nucleases, a formação de poros na célula antagônica, etc.

As bacteriocinas são proteínas ou peptídeos com baixo peso molecular que possuem a capacidade de inibir o crescimento bacteriano e estão presentes em bactérias Gram-positiva e Gram-negativas (COTTER; HILL; ROS, 2005; ATASSI; SERVIN, 2010;

DIEP, et al., 2009). São divididas em três classes: I (lantibióticos), II (bacteriocinas não modificadas) e III (proteínas termossensíveis) (DRIDER et al., 2006). Chatzidaki-Livanis *et al.* (2014) identificaram um gene que codifica a bacteriocina perforina em quatro linhagens de *B. fragilis* (638R, NCTC 9343, 379 eCL07T12C05). A perforina inibe o crescimento celular em outras linhagens de *B. fragilis*, mas não tem ação contra representantes de *Escherichia coli* e *Vibrio cholerae* (CHATZIDAKI-LIVANIS *et al.* 2014), indicando que podem ter um papel maior na competição intraespecífica.

1.5.3 Don PUL

Rogers *et al.* (2013) observou que *B. thetaiotaomicron* no ambiente colônico utiliza diferentes fontes de carbono, principalmente de polissacarídeos e glicanos provindos da dieta do hospedeiro. Essa adaptação deve-se a presença do *locus* de utilização de polissacarídeo (PULs) que foi descrito em *B. fragilis* por Cao, Rocha e Smith em 2014. Em seus estudos de expressão por microarranjo de DNA em modelo *in vivo* de abscesso em gaiola de tecido de rato, os autores demonstraram que o PUL em *B. fragilis* também aumenta a adaptação desta bactéria ao ambiente extraintestinal, onde a fonte de carbono é bastante limitada.

Cao, Rocha e Smith (2014) identificaram uma variedade de PULs, chamada Don PUL, que permite a retirada de glicanos de proteínas presentes no muco intestinal, fluido seroso do abscesso e soro (CAO; ROCHA; SMITH, 2014). A observação foi validada por RT-PCR (do inglês *reverse-transcriptase polymerase chain reaction* ou reação em cadeia da polimerase de transcriptase reversa). De acordo com este estudo, o *locus* Don PUL é composto por sete genes (*donA*, *donB*, *donC*, *donD*, *donE*, *donF* e *donG*). Os dois

primeiros atuam como reguladores, enquanto os demais codificam proteínas da membrana que possuem função de ligação e clivagem de glicanos das glicoproteínas.

1.5.4 Locus de Biossíntese de Polissacarídeo Capsular

O *Locus* de biossíntese de polissacarídeo capsular foi apontado como um importante fator de virulência em *B. fragilis* (ONDERDONK *et al.*, 1977). Ao longo do tempo, mais funções foram atribuídas aos genes deste *locus*. Ele é composto por 19 genes que regulam e expressam pelo menos três polissacarídeos capsulares, chamados polissacarídeo A (PSA), B (PSB) e C (PSC). Os polissacarídeos foram relacionados à formação de abscesso intra-abdominal e adesão (COYNE *et al.*, 2001); mas, como citado anteriormente também pode estimular a resposta imune adaptativa e também a inata. O polissacarídeo A pode desencadear forte resposta de células T CD4+, que por sua vez convoca células apresentadoras de antígeno portadoras de classe II (APC) e resulta em uma cascata de eventos imunológicos que culminam na formação do abscesso (COMSTOCK *et al.*, 1999; TZIANABOS *et al.*, 2000; COYNE *et al.*, 2001). Segundo Alvarez e colaboradores (2020), a resposta do hospedeiro ao PSA tem um efeito ambíguo, com a liberação de citocinas pró-inflamatórias e receptores de superfície anti-inflamatórios que podem causar ou inibir doenças inflamatórias intestinais.

1.5.5 Sistema Secretor Tipo VI (T6SS)

A sobrevivência microbiota colônica depende de adaptações que garantam maior vantagem na captação de recursos e na competição por espaço. Além de estratégias que envolvem o sistema CDI (inibição de crescimento dependente de contato) e a liberação de

toxinas e bacteriocinas, a competição intrainestinal também conta o Sistema Secretor Tipo VI (T6SS). Um sistema multiprotéico que causa a intoxicação de microrganismos antagônicos (CHATZIDAKI-LIVANIS; GEVA-ZATORSKY; COMSTOCK, 2016). A vantagem do T6SS é que ele não depende de receptores na célula-alvo para atingi-la (SONG *et al.*, 2021). Através de contato celular, este sistema transfere proteínas tóxicas (proteínas efetoras) a outros microrganismos (RUSSELL; PETERSON; MOUGOUS, 2014). A oclusão do sítio ativo impede a autointoxicação (BENZ; MEINHART, 2014; VERSTER *et al.*, 2017).

O sistema ajuda na colonização e enriquecimento de *Bacteroides* na microbiota intestinal (VERSTER *et al.*, 2017). Em Bacteroidales, o T6SS apresenta três arquiteturas: GA1, GA2 e GA3, sendo as duas primeiras bastante disseminadas dentro do filo; enquanto que GA3 é exclusiva de *B. fragilis* (CHATZIDAKI-LIVANIS; GEVA-ZATORSKY; COMSTOCK, 2016; COYNE; COMSTOCK, 2019).

O trabalho de Chatzidaki-Livanis, Geva-Zatorsky e Comstock (2016) descreveu como ocorre o funcionamento do T6SS em *B. fragilis*. De acordo com este estudo, o sistema possui proteínas efetoras e imunitárias únicas que são codificadas em duas regiões variáveis dentro do *locus*. O sistema envolve uma estrutura em forma de agulha localizada no citoplasma, similar à cauda do bacteriófago T4, e possui efetores tóxicos. Os efetores incluem enzimas que degradam NAD(P)⁺ (Fosfato de dinucleótido de nicotinamida e adenina), parede celular, membrana celular, fosfolipases, nucleases e toxinas que formam poros na célula antagonista. As proteínas imunitárias protegem a célula de suas proteínas efetoras e são codificadas por genes presentes em nas regiões variáveis. A Figura 5 apresenta a disposição dos genes ao longo do *locus* GA3 T6S.

menos desde o Pleistoceno. No entanto, esta resistência pode ser ancestral, visto que antibióticos derivadas de β -lactâmicos são produzidos na natureza por bactérias, fungos e insetos, como o colêmbolo *Folsomia candida* (CROFTS *et al.*, 2018; PONTAROTTI; RAOULT, 2019).

O *tetQ* (gene da proteína de proteção ribossômica resistente à tetraciclina) é um dos genes de resistência a antibióticos mais abundantes na microbiota intestinal humana de acordo com um estudo de metagenômica em populações da China, Dinamarca e Espanha (HU *et al.*, 2013). Os autores deste estudo apontaram que em 30 anos houve aumento de 50% na frequência deste gene, sugerindo a ocorrência de transferência horizontal no ambiente intestinal. O mecanismo de ação proporcionado pela proteína tetQ é a proteção física do sítio A ribossomal que impede a ação de tetraciclinas (CONNELL *et al.*, 2003). Niestępski e colaboradores (2019) encontraram este gene em 86.96% das linhagens de *B. fragilis*.

A presença do genes *erm* (metilase de resistência à eritromicina) confere resistência a macrolídeos, lincosamidas e estreptograminas, a chamada resistência MLS (PARK; KIM; JIN, 2010). Os antibióticos MLS apresentam um modo de ação similar que consiste na inibição da síntese proteica através da ligação à peptidil transferase do RNA ribossomal 23S (23S rRNA) da subunidade 50S do ribossomo bacteriano (BOZDOGAN; APPELBAUM, 2004). Quando uma metilase de resistência metila a adenina na posição 2058 do 23S rRNA a conformação do ribossomo é alterada e o antibiótico não se liga a seu alvo, tornando-se inócuo (MOAZED; NOLLER, 1987; BOZDOGAN; APPELBAUM, 2004). Dentre os genes *erm*, o *ermF* é o mais descrito em *B. fragilis* e está presente em 67.21% das linhagens analisadas por Niestępski *et al.* (2019).

1.6 LACTOBACILLACEAE

Lactobacillaceae são bactérias Gram-positivas, não esporulantes, catalase-negativas (a atividade da pseudocatalase pode estar presente em algumas linhagens), anaeróbicas facultativas, podem ser bastonetes ou cocobacilos não móveis e possuem baixo conteúdo G+C (FELIS; DELLAGLIO, 2007; MOZZI, 2016). A família é também chamada ‘Lactobacilli’ e inclui as bactérias do ácido láctico (LAB) que compartilham metabolismo e características fisiológicas na maioria relacionada à fermentação de alimentos e propriedades probióticas (GÄNZLE, 2015). Produzem ácido láctico a partir da fermentação de carboidratos (AGUIRRE; COLLINS, 1993).

São bastante heterogêneas e estão distribuídas em vários ambientes, como vegetais, laticínios, carne e cereais (fermentados), trato gastrointestinal de animais, boca e vagina, solo e água (GOLDSTEIN; TYRRELL; CITRON, 2015; MOZZI, 2016).

O gênero *Lactobacillus* foi proposto em 1901 por Beijerinck (*apud* Vandamme *et al.*, 1996) com base em características fenotípicas, utilização de carboidratos e metabólitos produzidos. Posteriormente outros critérios foram utilizados na classificação, como hibridização de DNA, conteúdo GC e estrutura química de peptidoglicanos (ZHENG *et al.*, 2020). A partir de 1983, com o trabalho de Schleifer e Stackebrandt, dados filogenéticos de 16S começaram a ser usados para embasar a classificação e nomenclatura deste gênero (SCHLEIFER; STACKEBRANDT, 1983). Além do 16S, a análise de fermentação de carboidratos com base na presença da enzima frutose-1,6-bisfosfato aldolase também foi considerada para embasar a taxonomia (VANDAMME *et al.*, 1996).

De acordo com esta proposta *Lactobacillus* é dividido em três grupos fisiológicos:

1) homofermentativo, que usa hexose para produzir ácido láctico; 2) heterofermentativo

facultativo, que utiliza tanto pentose quanto hexose para produção de ácido láctico e outros produtos; e 3) heterofermentativo obrigatório, que usa pentose ou hexose para produzir ácido láctico, gás carbônico e outros produtos e não possui a enzima frutose-1,6-bifosfato aldolase (VANDAMME *et al.*, 1996; SALVETTI *et al.*, 2018; ZHENG *et al.*, 2020).

O advento das análises genéticas nas últimas duas décadas revelou uma ampla diversidade no gênero, com isso várias espécies fossem incluídas e diversos subgrupos foram criados, mas com pouca resolução (SALVETTI; TORRIANI; FELIS, 2012; ZHENG *et al.*, 2015; ZHENG *et al.*, 2020).

As análises de genomas acabaram por revolucionar a classificação de *Lactobacillus* no último ano. Até março de 2020 o gênero *Lactobacillus* possuía 261 espécies, no entanto, uma revisão taxonômica baseada em genomas completos (identidade média entre nucleotídeos, aminoácidos e entre os genes do genoma central; filogenia do genoma central; genes de assinatura; e critérios metabólicos e ecológicos) realizada por Zheng e colaboradores (2020) redistribuiu estas espécies em 25 gêneros, que incluem o grupo *Lactobacillus delbrueckii*, *Lactiplantibacillus*, *Paralactobacillus*, e 22 outros gêneros (23 novos gêneros cujos nomes são *Holzapfelia*, *Amylolactobacillus*, *Bombilactobacillus*, *Companilactobacillus*, *Lapidilactobacillus*, *Agrilactobacillus*, *Schleiferilactobacillus*, *Loigolactobacillus*, *Lacticaseibacillus*, *Latilactobacillus*, *Dellaglia*, *Liquorilactobacillus*, *Ligilactobacillus*, *Furfurilactobacillus*, *Paucilactobacillus*, *Limosilactobacillus*, *Fructilactobacillus*, *Acetilactobacillus*, *Apilactobacillus*, *Levilactobacillus*, *Secundilactobacillus* e *Lentilactobacillus*). A proposta foi aceita pelo NCBI (*National Center for Biotechnology Information*; <https://www.ncbi.nlm.nih.gov/>) e LPSN (*List of Prokaryotic names with Standing in Nomenclature*; <https://www.bacterio.net>) (PARTE *et al.*, 2020). Os nomes das espécies adaptadas a animais foram mantidos. *Lactobacillus*

abrangendo espécies adaptadas a vertebrados (*L. acidophilus*, *L. apis* e *L. bombicola*, *L. delbrueckii*, *L. iners*, *L. crispatus*, *L. crispatus*, *jensensii* e *L. johnsonii*) e invertebrados (*Lactobacillus apis* e *Lactobacillus bombicola*) (ZHENG *et al.*, 2020).

Os 25 gêneros pertencem a Ordem Lactobacillales, Família Lactobacillaceae e são relacionados filogeneticamente aos gêneros *Paralactobacillus* e *Pediococcus* (GARRITY *et al.*, 2004; ZHENG *et al.*, 2020).

1.6.1 LAB e Probióticos

Um dos usos mais comuns das LAB é na produção de probióticos, que são microrganismos vivos ingeridos em determinadas quantidades e que proporcionam benefícios à saúde (GUARNER; SCHAAFSMA, 1998; TANNOCK, 2003; LJUNGH; WADSTRÖM, 2006). Além das LAB, outros microrganismos são usados para este fim, como *Bifidobacterium*, *Enterococcus*, *Saccharomyces*, *Bacillus* e *Clostridium butyricum* (KLEIN *et al.*, 1998; ELMER *et al.*, 1999; SENESI *et al.*, 2001; TAKAHASHI *et al.*, 2005).

Há relatos na literatura de pacientes com doenças específicas (colite ulcerativa, imunodeficiências, pacientes pediátricos em unidade intensiva, síndrome do intestino curto, cirurgia cardiovascular, etc) que tiveram piora no quadro clínico após o consumo de probióticos contendo *L. rhamnosus*, *L. acidophilus* e *L. casei* (VAHABNEZHAD *et al.*, 2013; MEINI *et al.*, 2015; LUONG *et al.*, 2010; HAGHIGHAT; CRUM-CIANFLON, 2016; CASTRO-GONZÁLEZ *et al.*, 2019). Além disso, as LAB podem possuir plasmídeos de resistência, geralmente de cloranfenicol, eritromicina e tetraciclina e outros

antibióticos (TANNOCK *et al.*, 1994; LIN *et al.*, 1996; GEVERS *et al.*, 2003; CASTRO-GONZÁLEZ *et al.*, 2019).

Mas, dado os cuidados necessários, os probióticos podem ajudar na modulação da microbiota intestinal causando benefícios à saúde. Por exemplo, *L. acidophilus* pode reduzir a inflamação intestinal e os pólipos causados por câncer no cólon, suprimindo a colite (PARK *et al.*, 2018; KHAZAIE *et al.*, 2012; CHEN *et al.*, 2015); *L. delbrueckii subsp. Bulgaricus* apresenta efeito antimutagênico em testes em animais (WOLLOWSKI; RECHKEMMER; POOL-ZOBEL, 2001); *L. fermentum* melhora o quadro clínico de pacientes com hipercolesterolemia (LEW; OHNO ; LIONG, 2017) e *Lactiplantibacillus plantarum* (previamente chamado *Lactobacillus plantarum*, mas reclassificado por Zheng *et al.* 2020) reduz os níveis de colesterol, triglicerídeo, o estresse oxidativo e também produz butirato, um AGCC que regula a homeostase intestinal (OOI; LIONG, 2010; ZHAO *et al.*, 2018; BOTTA *et al.*, 2017).

1.6.2 *Lactiplantibacillus plantarum*

L. plantarum é uma LAB heterofermentativa que ocupa diferentes nichos, como os tratos gastrointestinal, vaginal e urogenital; vegetais, produtos lácteos e alimentos fermentados (TORRIANI *et al.*, 2001; PFEILER; KLAENHAMMER, 2007; SIEZEN *et al.*, 2010; SIEZEN; VAN HYLCKAMA VLIEG, 2011; WANG *et al.*, 2011; NAMI *et al.*, 2014). Uma característica marcante destas bactérias é que conseguem obter energia de diferentes fontes de carboidratos e apresentam propriedades relevantes na manufatura de alimentos (embutidos, vinhos, probióticos) e na indústria farmacêutica (suplementos, produção de bacteriocinas, produtos anticâncer, antifúngicos e agentes com potencial

anticâncer) (SIEZEN *et al.*, 2010; SIEZEN *et al.*, 2011; TABASCO *et al.*, 2011; FRAS *et al.*, 2014; YANG; CHANG, 2010; LI; ZHOU; GU, 2016; AHN *et al.*, 2018).

L. plantatum é bastante utilizado no processamento de alimentos crus e possui duas importantes certificações de segurança alimentar de “Geralmente considerado seguro” (GRAS, do inglês *Generally recognized as safe*) e de “Presunção de segurança qualificada” (QPS, do inglês *Qualified Presumption of Safety*) (RAY; JOSHI, 2014; RICCI *et al.*, 2016).

Segundo Emser e colaboradores (2017), a procura por produtos minimamente processados vem crescendo a cada ano e *L. plantarum* apresenta propriedades desejáveis durante este processo. São bactérias seguras que conseguem sobreviver no trato gastrointestinal e aderem às células epiteliais, mantendo suas propriedades probióticas (JIA *et al.*, 2017). Além disso, muitas linhagens de *L. plantarum* apresentam resistência aos sais biliares e suco gástrico; tolerância à temperatura, pH e pressão dos ambientes estomacal e intestinais; atividade antibiótica (pela produção de peróxido de hidrogênio, bacteriocinas, ácido acético, ácido propiônico, ácido fenolático); susceptibilidade a antibióticos e habilidade imunomoduladora (SHOKRYAZDAN *et al.*, 2014; KALITA *et al.*, 2018; SAELIM; JAMPAPHAENG; MANEERAT, 2017; KWON *et al.*, 2017; BEHERA; RAY; ZDOLEC, 2018).

1.6.3 Propriedades Genômicas de *L. plantarum*

Em geral os genomas das LAB são reduzidos, mas os de *L. plantarum* estão entre os maiores, em média com 3 megabases (Mb) (SIEZEN; VAN HYLCKAMA VLIEG, 2011). Estes genomas apresentam muitos elementos móveis (prófagos, plasmídeos,

transposons e integrons) adquiridos via transferência horizontal de genes (HGT) (HUBERT; KAMMERER, 1994; KLEEREBEZEM *et al.*, 2003).

Uma importante região genômica adquirida por HGT foi descrita na linhagem WCFS1, é a chamada ilha de estilo de vida (*lifestyle islands*) que apresenta genes de regulação e que codificam proteínas de metabolismo e transporte de carboidratos (KLEEREBEZEM *et al.*, 2003; MOLENAAR *et al.*, 2005). A presença de muitos elementos móveis nos genomas alertam para a possibilidade de disseminação de genes de resistência ou determinantes de virulência (HOLDEN *et al.*, 2004; ALEKSHUN; LEVY, 2007; PARTRIDGE *et al.*, 2009; MUNOZ-LOPEZ; GARCIA-PEREZ, 2010; ZHANG; ZHANG; YE, 2011).

1.6.4 Bacteriocinas de *L. plantarum*

A produção de alimentos fermentados como carne, leite e vegetais é uma das tecnologias de processamento de alimentos mais antigas da história da humanidade. Seu primeiro registro data de aproximadamente seis mil anos antes de Cristo (a.C) no Oriente Médio (FOX, 1993). As LAB apresentam importante papel nesta preservação, pois são capazes de oxidar carboidratos gerando ácidos orgânicos, álcoois, dióxido de carbono como produtos finais (CAPLICE; FITZGERALD, 1999). Além dessas substâncias, as LAB também liberam compostos aromáticos, bacteriocinas, exopolissacarídeos e enzimas que além de preservar também auxiliam no aroma e sabor característicos de produtos fermentados (CAPLICE; FITZGERALD, 1999; LEROY; DE VUYST, 2004).

As bacteriocinas mais frequentemente encontradas em *L. plantarum* são nisina e plantaricinas. Não há informações sobre a estrutura da nisina produzida por *L. plantarum*. Em *Lactococcus lactis subsp. lactis*, este lantibiótico da Classe I possui 34 aminoácidos,

com peso molecular de 3,5 kDa e apresenta isoleucina e lisina como os aminoácidos terminais (BALCIUNAS *et al.*, 2013). No processamento pós-tradução da nisina são encontrados resíduos de aminoácidos incomuns como lantionina, β -metil lantionina e ainda resíduos desidratados (KLAENHAMMER, 1993; SAHL; BIERBAUM, 1998; TODOROV, 2009). São produzidas por bactérias Gram-positivas, incluindo as da ordem Lactobacillales (LUBELSKI *et al.*, 2008; DE ARAUZ *et al.*, 2009; SHIN *et al.*, 2016). São bastante utilizadas na conservação de alimentos por ter efeito de amplo espectro, estabilidade ao calor e tolerância a pH baixo (DELVES-BROUGHTON *et al.*, 1996; BONEV *et al.*, 2000). Seu uso foi considerado seguro a partir de 1969 de acordo com a Organização Conjunta para Alimentação e Agricultura/Organização Mundial da Saúde (FAO/OMS), mas apenas na década de 1980 foi aprovado nos Estados Unidos (1983 – FDA – do inglês US *Food and Drug Agency*) e Europa (1988 – aditivo alimentar número E234) (JOZALA; NOVAES; PESSOA JUNIOR, 2015).

As bacteriocinas da Classe II são divididas em IIa, IIb e IIc e possuem em comum a natureza não-lantibiótica (DRIDER *et al.*, 2006). Somente IIa, IIb estão presentes em *L. plantarum*. A pediocina PA-1 é uma bacteriocina de amplo espectro da Classe IIa que apresenta a sequência conservada (YGNGVXCXK/NXXC) em sua porção N-terminal e tem forte ação contra *Listeria monocytogenes*; enquanto que as plantaricinas da Classe IIb possuem a atividade condicionada à presença de dois peptídeos: EF ou JK (TURNER *et al.*, 1999; RODRÍGUEZ; MARTÍNEZ; KOK, 2002; NISSEN-MEYER *et al.*, 2009; DIEP *et al.*, 2009). Em algumas linhagens, as ORFs que codificam as plantaricinas estão organizadas em 5 operons: *plnJKLR*, *plnMNOP*, *plnABCD*, *plnEFI* e *plnGHSTUV* (DIEP; HÅVARSTEIN; NES, 1996). O operon *plnABCD* é responsável pela expressão da plantaricina A; enquanto que os operons *plnEFI* e *plnJKLR* codificam as plantaricinas EF e

JK e suas proteínas imunitárias, respectivamente; *plnGHSTUVW* codificam plnG e plnH e proteínas acessórias que juntos formam um sistema de transporte ABC que secretam peptídeos; *plnMNOP* codifica quatro proteínas putativas; e *plnXY* possuem função desconhecida (DIEP *et al.*, 2009). A organização destes genes é mostrada na Figura 6.

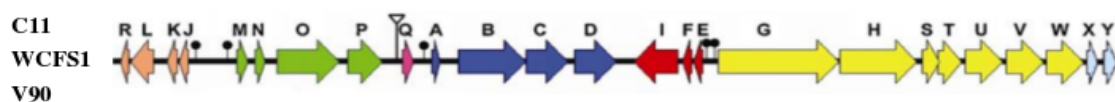


Figura 6: Modelo do *locus pln* das linhagens C11, WCFS11 e V90. A organização é similar, mas i) V90 possui um transposon simbolizado pelo triângulo, e *plnQ*, *plnS* e *plnT* aparecem como um único gene na imagem; iii) em C11, o gene *plnXY* está ausente na região à jusante de *plnW*. Os círculos indicam promotores regulados. Adaptado de Diep *et al.* (2009).

Além dessas, algumas linhagens de *L. plantarum* apresentam as plantaricina 163, BM-1, C, J51, S, ST8SH, ST28MS, W e Y (JIMENEZ-DIAZ *et al.*, 1995; ANDERSSSEN *et al.*, 1998; ZHANG *et al.*, 2013; SEDDIK *et al.*, 2017).

1.7 GENÔMICA COMPARATIVA

De acordo com os dados do *Bacterial Bioinformatics Resource Center* (PATRIC, <https://www.patricbrc.org/>), até dezembro de 2021 há disponíveis 464.574 sequências genômicas de Bactérias. 29.372 são genomas completos, os demais são sequências rascunho ou *draft* (*contigs* e *scaffolds*). Sequências genômicas apresentam informações de grande importância sobre as características genéticas dos organismos. Essas informações precisam ser interpretadas e comparadas a de outros táxons para estabelecer relações filogenéticas e identificar processos evolutivos. As diferentes ferramentas de Genômica

Comparativa permitem contrapor características genômicas simples como a composição G+C e uso de códon dos genomas; além de dados mais complexos, que permitem o alinhamento genômico, predições gênica e funcional, identificação de elementos genéticos móveis e sequências CRISPR-Cas, análise da evolução molecular e etc. Os tópicos a seguir descrevem as diferentes ferramentas usadas para a análise de genômica comparativa. Algumas delas foram usadas nos dois artigos que compõem esta tese.

1.7.1 Alinhamento Genômico

Os genomas de organismos filogeneticamente próximos tendem a compartilhar características genômicas e blocos sintênicos, que são regiões que possuem a mesma ordem de genes homólogos derivada de um ancestral comum (VERGARA; CHEN, 2010; TANG *et al.*, 2011). Esses blocos, assim como outras características genômicas, podem ser informativos durante a predição gênica (DE MAN *et al.*, 2016; HUNT *et al.*, 2016). Várias ferramentas foram desenvolvidas para comparar genomas, algumas delas são o ACT, MAUVE e RING.

O ACT (Artemis Comparison Tool-ACT) é uma aplicação em JAVA que permite a comparação de qualquer tipo de genoma (circular ou linear) e aceita arquivos de entrada em diferentes formatos (EMBL, Genbank, GFF e FASTA), ou um arquivo com as comparações entre os genomas gerados pelos programas BLASTN (*Basic Local Alignment Search Tool* – para análise de nucleotídeos), BLASTX (*Basic Local Alignment Search Tool* – para análise de proteína) ou MUMmer (CAMACHO *et al.*, 2009; CARVER *et al.*, 2005). O download é disponível em: <http://sanger-pathogens.github.io/Artemis/ACT/> ; CARVER *et al.*, 2005).

O RING (*BLAST Ring Image Generator*) é uma multiplataforma escrita na linguagem JAVA que utiliza os programas CGView (TOTHARD; ISHART, 2005) e BLAST (CAMACHO *et al.*, 2009) para comparar genomas completos. Os resultados são plotadas em imagens em círculos que correspondem a cada um dos genomas (Download disponível em: <http://brig.sourceforge.net/> ; ALIKHAN *et al.*, 2011).

O MAUVE (*Multiple genome alignment*), é um programa em JAVA que permite alinhar e identificar a presença de rearranjos, inserções ou deleções de genes ou conjuntos de genes de genomas bacterianos (DARLING *et al.*, 2004). O download é pelo link: <http://darlinglab.org/mauve/mauve.html>.

1.7.2 Pan-genoma

O avanço do sequenciamento de genomas bacterianos mostrou que cada táxon possui genes essenciais ao metabolismo e à fisiologia celular e também há genes dispensáveis. A partir da descrição destas categorias foi cunhado o termo Pan-genoma (pan do grego ‘todo’), que é a soma de todos os genes de um genoma (TETTELIN *et al.*, 2005; MEDINI *et al.*, 2005). Os genes essenciais compõem o genoma central e estão presentes em todos os representantes do táxon, como exemplos são os genes ribossomais, do envelope celular, replicação de DNA e metabolismo; o pool gênico flexível ou dispensável possui genes acessórios, que estão somente em algumas linhagens; e há os chamados genes únicos, bastante restritos em sua maioria não caracterizados (READ; USSERY, 2006; MIRA *et al.*, 2010).

Durante a análise de pan-genoma, à medida que novas sequências genômicas são adicionadas à pesquisa, o conjunto de genes tende a aumentar, principalmente em bactérias

adaptadas a diferentes ambientes e com ampla troca de material genético (VERNIKOS, 2020). Neste caso, o pan-genoma é dito aberto. As bactérias que vivem em nichos ecológicos mais restritos tendem a receber menos genes externos de outras bactérias e seus pan-genomas tendem a ser fechados (VERNIKOS, 2020). Com base nessas informações é possível saber se mais linhagens deverão ter os genomas sequenciados ou não. Ou seja, uma bactéria com pan-genoma fechado já teve seu pool gênico revelado e geralmente pouca novidade será encontrada em novas pesquisas genômicas; enquanto que bactérias com pan-genomas abertos precisam ser melhor representadas para maior compreensão de sua composição gênica.

A análise pode gerar outras informações, como uma inferência filogenética mais completa, como a utilizada por Zheng e colaboradores (2020) no reposicionamento taxonômico do gênero *Lactobacillus* (indicado posteriormente).

Algumas ferramentas usadas para a análise dos pan-genomas são PGAP, Roary e *Bacterial Pangenome Analysis* (BPGA) (ZHAO *et al.*, 2012; ANDREW *et al.*, 2015; CHAUDHARI *et al.* 2016).

O PGAP - *Pan-Genomes Analysis Pipeline* utiliza a Perl e usa vários métodos para realizar a análise genômica (ZHAO *et al.*, 2012). Depende da instalação de outros programas como o BLAST, mcl, PHYLIP, etc. Cada amostra precisa de três arquivos de entrada com as informações das proteínas, nucleotídeos e anotação do genoma (.pep, .nuc, .function). Os arquivos de saída são na extensão .txt e os dados devem ser convertidos por outros programas (também escritos em Perl) para a interpretação dos resultados. O download é pelo link <https://sourceforge.net/projects/pgap/>.

O Roary, um programa que pode ser baixado pelo link <https://sanger-pathogens.github.io/Roary/>, é escrito na linguagem Perl e depende de várias dependências durante a sua instalação (como cd-hit, ncbi-blast+, fasttree, etc.). O arquivo de entrada é no formato *General Feature Format* (GFF3) gerado pelo Prokka a partir de um arquivo genbank (SEEMANN, 2014). O programa gera vários arquivos de saída, conforme a configuração. Os arquivos gerados podem ser visualizados em plataformas como Phandango (<http://jameshadfield.github.io/phandango/#/>) e panX (<http://pangenome.tuebingen.mpg.de/#contacts>) (HADFIELD *et al.*, 2018; DING; BAUMDICKER; NEHER, 2018).

O BPGA também é escrito em Perl e analisa arquivos de entrada em três formatos (gbk, faa e pep.fsa). As dependências do programa auxiliam na criação do arquivo de alinhamento múltiplo (MUSCLE - *MUltiple Sequence Comparison by Log- Expectation*), os gráficos (Gnuplot 4.6.6) e também no cruzamento dos resultados com dados funcionais dos bancos de dados COG (*Clusters of Orthologous Groups of proteins*) e do KEGG (*Kyoto Encyclopedia of Genes and Genomes*) (USEARCH, CD-HIT e OrthoMCL) (EDGAR, 2004; EDGAR, 2010; WEIZHONG *et al.*, 2012; LI; STOECKERT; ROOS, 2003; CHAUDHARI *et al.* 2016). O download é pelo link <https://iicb.res.in/bpga/>.

1.7.3 Sequências CRISPR-Cas

As repetições palindrômicas curtas agrupadas e regularmente interespacia (CRISPR) e suas proteínas associadas (Cas) formam o chamado sistema CRISPR-Cas (HORVATH; BARRANGOU, 2010). As sequências CRISPR são compostas por regiões repetidas altamente conservadas (DR) com comprimento de 24 a 50 bp que são separadas

por espaçadores, aparentemente fragmentos de vírus ou plasmídeos (COUVIN *et al.*, 2018). Estão presentes em bactérias e arqueas e fazem parte do sistema imune contra a invasão de DNA exógeno (BARRANGOU; RODOLPHE; CHRISTOPHE, 2007). O sistema possui uma constituição básica dividido em dois módulos: o módulo adaptativo que possui genes para gerar RNA CRISPR (crRNA); e o módulo efetor com genes regulados pelo crRNA e com função de clivar o material genético invasor (MAKAROVA *et al.*, 2011). O sistema CRISPR é classificado em duas classes, seis tipos e 22 subtipos que variam de acordo com a arquitetura do *loci* genômico (MAKAROVA *et al.*, 2015). A classe I compreende os tipos I, III e IV; e a Classe 2 os tipos II e V. Ambas possuem em comum as proteínas Cas1 e Cas2, mas a Classe I possui complexos efetores de várias subunidades compostos por múltiplas proteínas Cas; enquanto que da Classe II possui um complexo efetor com uma longa e única proteína Cas (SHMAKOV *et al.*, 2015).

Algumas ferramentas para detecção das sequências CRISPR são CRT, PILER-CR e CRISPRCasFinder. A CRT (*CRISPR Recognition Tool*) é uma ferramenta escrita em Java que detecta diretamente as repetições DR de sequências de DNA por isso bastante rápida em análises de genomas e metagenomas (BLAND *et al.*, 2007). O programa está disponível em <http://www.room220.com/crt/>.

PILER-CR é escrita na linguagem C e tem a proposta de identificar e classificar as repetições DR de forma rápida (cerca de 5 Mb em cinco segundo). O programa possui um catálogo das repetições de genomas de 346 bactérias e arqueias que agiliza as buscas (EDGAR, 2007, disponível em <http://www.drive5.com/pilercr>).

O CRISPRCasFinder é uma atualização dos programas CRISPRFinder e CasFinder, sendo o arquivo de entrada no formato fasta (.fas) (COUVIN *et al.*, 2018). Esta versão permite a identificação dos tipos da sequência CRISPR e a detecção das proteínas

Cas. Além da interface Web é possível baixar o programa que é escrito em Perl e é compatível com todos os sistemas operacionais (COUVIN *et al.*, 2018; Disponível em <https://crisprcas.i2bc.paris-saclay.fr>).

1.7.4 Predição do Mobiloma

Cerca de 75% dos genes dos genomas bacterianos foram adquiridos por transferência horizontal de genes (HGT, do inglês *horizontal gene transfer*) (KLOESGES *et al.*, 2011). Portanto, ela tem um grande impacto na plasticidade e adaptação e é promovida por elementos genéticos móveis (ou MGE - *mobile genetic element*), como elementos de inserção (IS), elementos integrativos e conjugativos (ICEs) e elementos integrativos e mobilizáveis (IMEs), plasmídeos, transposons (Tn), transposons conjugativos (Ctn), prófagos e ilhas genômicas (JUHAS *et al.*, 2013; CARR *et al.*, 2021).

Os ICEs e IMEs foram descobertos através de análises de bioinformática. De acordo com Johnson e Grossman (2015), as ICE putativas com T4SS, que são ICEs associados à maquinária proteica que envolve o sistema de secreção bacteriana do tipo IV (T4SS), proteína de acoplamento tipo IV (T4CP) e relaxase, possuem grande papel na disseminação de genes de resistência a antibióticos em bactérias gram-negativas, pois são capazes de conjugação e mobilização nas formas cis e trans (CHRISTIE *et al.*, 2014; ORLEK *et al.*, 2017; CHE *et al.*, 2021). ICEs e IMEs podem apresentar os sítios de recombinação *attL* e *attR* (sequências DR – *direct repeat*) (LIU *et al.*, 2018).

A detecção de elementos genéticos móveis é de ampla importância para entender a dinâmica evolutiva bacteriana. Muitas toxinas são codificadas por genes de bacteriófagos, como exemplo as toxinas do *Vibrio cholerae*, *Corynebacterium diphtheriae* e *Clostridium difficile* (PHAM *et al.*, 2018; HOLMES, 2000; FORTIER, 2017; KASMAN;

PORTER, 2020). Além disso, bacteriófagos também podem transportar genes de resistência a antibióticos e outros determinantes de virulência (BOYD, 2012; KASMAN; PORTER, 2020). Alguns dos programas usados para identificação de bacteriófagos nos genomas bacterianos são o PHASTER e PhiSpy.

O PHASTER (*PHAge Search Tool Enhanced Release*) foi originalmente escrito na linguagem Perl e apresenta a interface web com base em CGI (interface comum de ligação), mas foi reescrito na multiplataforma de código aberto *Ruby on Rails*. As buscas no genoma bacteriano por regiões profágicas contendo genes com funções conhecidas como terminases, lisinas, terminases e transposases são através do BLASTP (ARNDT *et al.*, 2016). Os arquivos de entrada são nos formatos fasta (.fas) e genbank (.gbk) e os resultados mostram as supostas regiões do prófago, sua localização no genoma bacteriano, os supostos fagos de origem, conteúdo GC e proteínas fágicas.

O PhiSpy é um algoritmo escrito nas linguagens Python e C++ e considera o enviesamento GC, comprimento da proteína, orientação do transcrito, abundância da palavra fago, homologia de 40 genes conhecidos para fazer suas buscas. O arquivo de entrada é o genbank (.gb) (AKHTER; AZIZ ; EDWARDS, 2012). O programa está disponível para download pelo link <https://github.com/linsalrob/PhiSpy>.

Os MGE mais simples são os IS, que possuem comprimentos que variam de 700 a 2.500 pb e utilizam apenas a proteína transposase para fazer a sua mobilização ao longo do genoma (KUSUMOTO *et al.*, 2011). A presença desses elementos pode ser detectada através do uso do MobileElementFinder (V1.0.3) (LIU *et al.*, 2019; JOHANSSON *et al.*, 2021), disponível em (<https://cge.cbs.dtu.dk/services/MobileElementFinder/>). Este programa é uma atualização de uma versão de 2019 e foi desenvolvido pelo Centro de genômica Epidemiológica na Dinamarca. Ele foi escrito em *Python* v3.7 e pode ser usado

na interface web ou instalado (JOHANSSON *et al.*, 2021). Os IS são curtos, portanto, não há obrigatoriedade em analisá-los somente em genomas completos; por outro lado, o ICEFinder, detecta ICEs e IMEs (do inglês, *integrative and mobilizable elements*), que possuem entre 5 a 50 Kb (GUÉDON *et al.*, 2017). Sendo mais apropriado seu uso em genomas completos (JOHANSSON *et al.*, 2021). Os dois programas indicam se o MGE carrega genes de resistência ou determinante de virulência e a informação pode ser verificada manualmente no genoma.

1.7.5 Genes de Resistência a Antibióticos

A resistência aos antibióticos por bactérias Gram-negativas constitui uma ameaça à saúde pública que foi agravada desde 1970 e foi acentuada pela disseminação de genes de resistência via HGT (FALAGAS; KASIAKOU; COLISTIN, 2005; DOMINGUES *et al.*, 2012). Até o momento, há poucas alternativas para o tratamento de patógenos específicos, principalmente em infecções adquiridas no ambiente hospitalar (SUN *et al.*, 2018; AHMED *et al.*, 2020). Diferentes condições de resistência encontradas em bactérias chamadas de multirresistência a medicamentos, extensivamente resistente a medicamentos e linhagens resistentes a pandrogas foram descritas em linhagens de *E. coli*, *Klebsiella pneumoniae*, *Acinetobacter baumannii* e *Pseudomonas aeruginosa* no mundo inteiro (TZOUVELEKIS *et al.*, 2012; MUNOZ-PRICE *et al.*, 2013; KEMPF *et al.*, 2012; JEANNOT; BOLARD; PLÉSIAT, 2016). Estas bactérias são encontradas na microbiota intestinal e podem disseminar facilmente genes de resistência via HGT.

A identificação de genes de resistência nos genomas bacterianos é de ampla importância e algumas ferramentas fazem essa predição. Algumas delas são o Arg-annot, CARD e ResFinder 4.0.

Arg-annot (Antibióticos Resistência Gene-ANOTAÇÃO) é uma ferramenta que utiliza o BLAST local (BLASTN, BLASTP ou BLASTX) do BioEdit para analisar sequências. A ferramenta usa um banco de dados baseado em 1689 genes de resistência a antibióticos coletados da literatura e de outros bancos de dados (GUPTA *et al.*, 2014). Pode executar uma sequência ou múltiplas sequências no formato fasta (.fas). As análises podem ser realizadas pelo link https://ifr48.timone.univ-mrs.fr/blast/arg-annot_nt.html

O CARD (Banco de dados abrangente de resistência a antibióticos) (ALCOCK *et al.* 2020, disponível em <https://card.mcmaster.ca>) integra biologia molecular, bioquímica e bioinformática em uma estrutura para produzir um banco de dados que é funcional e prático. O esquema e os dados são gerenciados com o PostgreSQL 9.5. Além deste serviço de banco de dados, o site e ferramentas são projetados com o Laravel 5.2 PHP, PHP 7.0.22 e Apache 2.4 e as estatísticas complementares são geradas com o Biopython (ALCOCK *et al.* 2020). Os dados selecionados pelo CARD são organizados obedecendo três passos: 1) O ARO, que contém a base primária das buscas por genes de resistência a antibióticos, alvos de drogas, classe de drogas e mecanismos moleculares de resistência; 2) o modelo CARD (MO); e 3) NCBITaxon, um subconjunto com curadoria pelo Ontologia de taxonomia organizacional do NCBI. As buscas por genes de resistências podem ser realizadas pelos algoritmos BLAST ou RGI (identificador do gene da resistência). Em ambos, o arquivo de entrada é uma sequência fasta (.fas). No RGI a sequência pode corresponder a um genoma completo ou *draft* e o resultado será a predição de todos os prováveis genes de resistência a antibiótico presentes no genoma; no BLAST uma sequência gênica candidata poderá ser inserida e as estatísticas sobre o gene serão geradas para posterior análise.

O ResFinder 4.0 foi reescrito em Python 3, enquanto as versões anteriores foram escritas em Perl. A nova versão permite o alinhamento das sequências alvo às de bancos de dados sem a necessidade de montagem, estratégia que agilizou o programa. As análises podem ser realizadas diretamente pelo servidor (disponível em <https://cge.cbs.dtu.dk/services/ResFinder/>) ou o programa pode ser instalado diretamente pelo usuário (disponível em: <https://bitbucket.org/genomicepidemiology/resfinder/src/master/>) (BORTOLAIA *et al.*, 2020).

1.7.6 Genes de Virulência

A patogenicidade de um microorganismo é acentuada pela atuação dos fatores de virulência, que são moléculas que aumentam a habilidade do parasita auxiliando na sua disseminação, infecção e fuga do sistema imune hospedeiro, tendo como exemplo de fatores de virulência as toxinas, cápsula, adesinas, etc (CROSS, 2008). Vários bancos de dados com diferentes fatores de virulência estão disponíveis na internet, a maioria deles possuem dados integrados e curados continuamente. O banco de dados do fator de virulência VFDB (*The virulence factor database*, <http://www.mgc.ac.cn/VFs/main.htm>) possui informações gerais sobre as principais bactérias patogênicas e ainda disponibiliza busca por genes candidatos a fatores de virulência através do tradicional BLAST ou pela ferramenta Vfanalyzer (LIU *et al.*, 2019). O VFs ou Victors (<http://www.phidias.us/victors>) é outra plataforma com conteúdo amplo e vários fatores podem ser combinados durante as buscas nome dos patógenos (bactéria, fungos e vírus), informações gênicas e funcionais (COG) sobre o gene de interesse (SAYERS *et al.*, 2019). O PATRIC é uma multiplataforma de bioinformática que também integra dados do

PATRIC_VF (banco de dados interno), VFDB e Victor's para a busca de fatores de virulência em genomas ou táxons específicos (DAVIS *et al.*, 2020).

1.7.7 Análise de Evolução Molecular

A seleção positiva é o processo onde uma nova mutação ou alelo raro anteriormente consegue aumentar a aptidão de um indivíduo (NIELSEN *et al.*, 2007; VITTI *et al.*, 2013; LÓPEZ; NEIRA; YÁÑEZ, 2015). Enquanto que a seleção purificadora ou negativa é quando substituições ou alelos novos são deletados (VITTI *et al.*, 2013). Tradicionalmente a pressão seletiva, simbolizada pela letra grega ômega (ω) é estimada pela razão dN/dS (YANG, 2003). Onde dN corresponde às substituições não-sinônimas e dS são as substituições sinônimas em uma sequência proteica. O valor de ω médio, considera três situações possíveis: 1) dN= dS ($\omega = 1$, seleção neutra); 2) dN < dS ($\omega < 1$, seleção purificadora); 3) dN > dS ($\omega > 1$, seleção positiva ou adaptativa) (YANG, 2003). Esse tipo de análise é importante para testar a hipótese de neutralidade sobre uma sequência e por isso ainda é bastante usada em testes e programas de análises filogenéticas e populacionais, como o Arlequin e MEGA (*Molecular Evolutionary Genetics Analysis*) (EXCOFFIER; LISCHER, 2010; KUMAR *et al.*, 2020). No entanto, essa abordagem é pouco realista e falha em indicar seleção positiva em poucos códons, como por exemplo nos genes da protease do HIV-1 e Spike (D614G) do Covid-19 (CRANDALL *et al.*, 1999; KORBER *et al.*, 2020).

Em 1991, Gillespie observou que a evolução adaptativa é restrita a alguns momentos da história evolutiva (também chamada seleção darwiniana episódica) e afeta poucos códons e não a sequência completa, visto que a função proteica poderia ser comprometida. Além disso, Messier e Stewart (1997) consideraram em suas análises que o

valor de ω varia de acordo com a relação filogenética entre os táxons analisados e as proteínas possuem domínios mais conservados que outros, portanto valores de ω podem diferir ao longo de seus sítios. Posteriormente, essas observações foram consideradas e modelos de verossimilhança mais realistas foram criados (YANG, 1998; YANG; NIELSEN, 1998; NIELSEN; YANG, 1998; YANG, 2000).

O desenvolvimento de novas metodologias aumentou a compreensão sobre o mecanismo de evolução molecular e vários programas utilizam a nova abordagem para a busca de táxons ou sítios sob seleção positiva. Uma das primeiras ferramentas a executar este tipo de análise foi o PAML (*Phylogenetic Analysis by Maximum Likelihood*), um pacote de programas de análise de evolução molecular escrito em C++ que teve alterações ao longo dos anos. Sua versão anterior era PAML 4 (YANG, 2007), mas em 2013 a versão PAMLX ganhou um interface gráfica que pode ser executada em todos os sistemas operacionais (XU; YANG, 2013). O pacote está disponível em <http://abacus.gene.ucl.ac.uk/software/paml.html>. Ele considera duas opções do modelo dos ramos que permite que o valor de ω varie em uma filogenia (YANG; NIELSEN 1998). O segundo modelo é o modelo dos sítios, onde o valor de ω pode variar entre os códons da sequência (YANG *et al.*, 2000).

O Hyphy (*Hypothesis Testing using Phylogenies*) é um pacote de programas que utiliza uma linguagem similar ao C (HyPhyBatch Language ou HBL) e GUI (KOSAKOVSKY POND *et al.*, 2020). E assim como o PAML utiliza métodos baseados em verossimilhança para buscar os sítios sob seleção, mas também pode usar a inferência bayesiana em alguns modelos (como exemplo o FUBAR). O pacote tem duas opções de execução, diretamente pelo servidor Datamonkey (disponível em <https://www.datamonkey.org>) ou via linha de comando no computador do usuário,

baixando e instalando o pacote Hyphy 2.5 (disponível em <https://www.hyphy.org/register/>). Na primeira opção somente o arquivo de entrada no formato nexus (.nex) precisa ser fornecido; enquanto que na execução na máquina do usuário é necessário ainda adicionar uma árvore filogenética no formato newick. Os arquivos de saída gerados pelo Hyphy são mais completos do que àqueles do Datamonkey, de forma que a instalação do programa é uma opção mais confiável. Tanto o Hyphy quanto o Datamonkey tiveram atualizações, mas ambos possuem as ferramentas GARD (*Genetic Algorithm Recombination Detection*), para a análise de recombinação, além de FEL (*Random effects likelihood*), FUBAR (*Fast Unconstrained Bayesian Approximation*) e SLAC (*Single-likelihood ancestor counting*) para análise de pressão seletiva. O modelo evolutivo MG94xREV é o padrão para rodar as análises. Ele aprimora os comprimentos dos ramos nas filogenias e as substituições de nucleotídeos das sequências (KOSAKOVSKY POND *et al.*, 2006; KOSAKOVSKY POND, 2005; MURRELL *et al.*, 2013; KOSAKOVSKY POND *et al.*, 2020).

1.8 OBJETIVOS

1.8.1 Objetivo Geral

Analisar os genomas de duas bactérias da microbiota intestinal humana, *B. fragilis* e *L. plantarum* usando diferentes ferramentas de genômica comparativa.

1.8.2 Objetivos Específicos

- Descrever os pan-genomas das bactérias;
- Inferir as sequências CRISPR-Cas dos genomas;
- Identificar possíveis determinantes de virulência;
- Verificar a presença de genes de resistência a antibióticos;
- Pesquisar a evolução molecular de *B. Fragilis*.

2. CAPÍTULO I

COMPARATIVE GENOMICS ANALYSIS OF *BACTEROIDES FRAGILIS*

Artigo submetido à revista: Genetics and Molecular Biology

Genetics and Molecular Biology



Comparative genomics analysis of *Bacteroides fragilis*

Journal:	<i>Genetics and Molecular Biology</i>
Manuscript ID	GMB-2021-0347
Manuscript Type:	Original Article
Date Submitted by the Author:	27-Oct-2021
Complete List of Authors:	Evanovich, Eliane; Universidade Federal do Pará, Instituto de Ciências Biológicas Mendonça-Mattos, Patricia ; Universidade Federal do Pará, Instituto de Ciências Biológicas ; Fundação Centro de Hemoterapia, Fundação Centro de Hemoterapia e Hematologia do Pará Guerreiro, João ; Universidade Federal do Pará, Instituto de Ciências Biológicas
Keyword:	<i>Bacteroides fragilis</i> , Comparative genomics, Genes for antimicrobial resistance, Horizontal gene transfer

SCHOLARONE™
Manuscripts

Comparative genomics analysis of *Bacteroides fragilis*

Short running: Genomic in *Bacteroides fragilis*

Eliane Evanovich*¹, Patricia Jeanne de Souza Mendonça-Mattos^{1,2} João Farias Guerreiro¹

Filiation:

¹Laboratório de Genética Humana e Médica, Instituto de Ciências Biológicas - Universidade

Federal do Pará, Pará, Brasil.

²Laboratório de Imunogenética, Fundação Centro de Hemoterapia e Hematologia do Pará, Pará, Brasil.

Address for correspondence:

E-mail: lianevanovich@gmail.com

Mailing address: Laboratório de Genética Humana e Médica - Universidade Federal do Pará, Rua Augusto Corrêa, 1 - Guamá, Belém - PA, 66075-110.

ORCID ID: <https://orcid.org/0000-0001-7076-5621>

Abstract

Bacteroides fragilis are commensal bacteria of the gut microbiota of mammals and may cause severe infection in a susceptible host. Treatment can be cumbersome if multidrug-resistant strains are present in the affected tissue. The principal aim of this study was to provide new insights into the genomic properties of *B. fragilis* through different approaches in comparative genomics. Results revealed that the pan-genome is opened, and an intense exchange of genetic material reinforces this inference. Don complex, responsible for extraintestinal adaptation, is present in all strains, suggesting a crucial role for *B. fragilis* adaptation. CRISPR-Cas system is in 76% of samples, but it seems to have low accuracy against prophages. Multidrug-resistance genes are in 80% of strains. Conjugative transposons and integrative and conjugative elements (ICE) are the main spreaders of genes for antimicrobial resistance. We also reported evidence for horizontal gene transfer (HGT) of antimicrobial resistance genes among the *B. fragilis* strains and within Order Bacteroidales. At least 398 genes are under positive selection, including genes for antimicrobial resistance and transport of toxins and nutrients.

Keywords: *Bacteroides fragilis*, Comparative genomics, Genes for antimicrobial resistance, Horizontal gene transfer.

1. Introduction

Bacteroides fragilis is an obligatory anaerobic Gram-negative bacillus. It is usually a commensal in mammals and is critical for immune development and intestinal mucosal integrity (Finegold, 1995; Round and Mazmanian, 2010; Mazmanian et al., 2008). It may become pathogenic in a vulnerable host, being able to invade intra- and extra-abdominal tissues, including blood (Sears et al., 1995; Wexler, 2007; Anka et al., 2015; Papaparaskevas et al., 2011; Valguarnera and Wardenburg, 2020; Elsaghir and Reddivari, 2021). Recent evidence suggests that enterotoxigenic *B. fragilis* (ETBF) is associated with colorectal cancer development (Vacante et al., 2020; Cheng et al., 2020; Valguarnera and Wardenburg, 2020; Elsaghir and Reddivari, 2021).

ETBF can produce fragilysin (known as BFT), an endotoxin from the zinc-dependent metalloproteinase family that cleaves E-cadherin of colonocytes (Moncrief et al., 1998; Franco, 2004; Shiryayev et al., 2014; Herrou et al., 2016; Bamias and Cominelli, 2016). The interaction between BFT and intestinal epithelial cells triggers an inflammatory response that increases apoptosis and favours extraintestinal invasion (Ko et al., 2016). *BFTP* gene encodes the BFT and has three alleles: *BFTP-1*, *BFTP-2* and *BFTP-3* (Franco et al., 1997; Kato et al., 2000; Sears, 2001). It is on a pathogenic island termed BfPAI, along with genes for a hypothetical protein and metalloproteinase II (*MPII*) (Franco et al., 1997; Kato et al., 2000; Sears, 2001; Franco et al., 2002; Franco, 2004; Shiryayev et al., 2014). BfPAI of some ETBFs is in the CTn86 transposon (Franco, 2004; Shiryayev et al., 2014).

Antimicrobial multidrug-resistance (AMR) genes are widespread among the *B. fragilis* strains and hamper antimicrobial therapy (Niestępski, 2017). Integrative and conjugative elements (ICE) disseminate AMR genes because they can mobilize plasmids, genomic islands, and other non-conjugative elements (Johnson and Grossman, 2015). In addition, their genes encode proteins for excision, conjugation (type IV secretion system-T4SS), and transferring DNA sequences (Guglielmini et al., 2011).

Other genes that improve *B. fragilis* adaptation to intestinal and extraintestinal environments belong to Type VI secretion systems (T6SS), capsular lipopolysaccharide biosynthesis locus (LPS locus), and Don PUL complex. T6SS is a protein complex that acts during the competition, inhibiting or killing other bacteria or eukaryotic cells (Russell et al., 2011; Ho et al., 2014; Chatzidaki-Livanis et al., 2016). LPS locus has 19 orthologous

genes (OG) that control the expression of three capsular polysaccharides (polysaccharide A, B, and C) and may trigger the formation of intra-abdominal abscesses (Adegbola, 1998; Onderdonk et al., 1977; Coyne, 2001; Comstock, 1999; Vestby et al., 2020). Don PUL complex enables extraintestinal adaptation through uptake of N-linked glycans from fluids (Cao et al., 2014).

Despite the approaches described above, few studies have investigated the comparative genomics of *B. fragilis*. Thus, this paper aims to provide insights into the genomic properties of these bacteria. For this purpose, we analysed 183 strains applying comparative genomics tools. Some insights obtained were: 1) Most genes are related to metabolic processes, mainly carbohydrates and amino acids - a frequent pattern of bacterial gut microbiota; 2) The Don PUL is the only system common to all strains, suggesting that alternative resource capture is decisive for their survival in the gut microbiota; and 3) We also reported evidence of HGT within order Bacteroidales. Hence, this research extends our knowledge of the genomic evolution of *B. fragilis*.

2. Materials and Methods

Sample

Samples used were available at the National Center for Biotechnology Information - NCBI (<https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/414/>). Genome lengths range from 4.87 to 7.60 megabases (complete, scaffold, and contig), and minimal coverage was 15X (Sims et al., 2014). The strains used and their sources are in the additional file, Table S1.

Pan-genome Analysis

Bacterial Pan Genome Analysis tool version 1.3 (BPGA) was used to identify core, accessory, and unique genomes (Chaudhari et al., 2016). Gene functions were based on Clusters of Orthologous Groups of proteins (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG). Orthologous cluster analysis (default setting) was performed by Usearch9.2.64 (Edgar, 2004). Multiple sequence comparison by log-expectation (MUSCLE) was used to generate alignments and phylogenies (Edgar, 2004). Gnuplot 5.0 was used for plotting results into graphs (Williams and Kelley, 2018).

Prediction of CRISPR and Resistance genes

CRISPR-Cas sequences were estimated by CRISPRRCasFinder (<https://crisprcas.i2bc.paris-saclay.fr/CrisprCasFinder/Index>) (Couvin et al., 2018). Antimicrobial resistance genes were obtained by CARD (Comprehensive Antibiotic Resistance Database) (<http://arpcard.mcmaster.ca/>). CARD was performed under the Resistance Gene Identifier option - RGI (Perfect match equal to 100) (Jia et al., 2017). Results with a 'strict match' between 95.0 and 99.9 also were analyzed under the Basic Local Alignment Search Tool- BLAST, as suggested by Jia et al. (2017).

Prediction of Mobilome

PhiSpy software was used to predict putative prophages, and ICEFinder was performed to identify ICE and IME (Integrative mobilizable elements) (Akhter et al., 2012; Liu et al., 2019). These analyses were restricted to 19 complete genomes, because prophage, ICE and IME regions are usually long and would be cut into short sequences, such as contigs and scaffolds.

Candidate proteins were searched against the CARD (BLASTp) (Alcock et al., 2020), Virulence factor database (VFDB) [http://www.mgc.ac.cn/VFs/search_VFs.htm; blastp/Protein sequences from VFDB full dataset (setB)] (Chen et al., 2005), and NCBI server (BLASTp) (<https://blast.ncbi.nlm.nih.gov/>) (Altschul et al., 1997). The score values significant were bit-score equal to or higher than 50; E-value ≤ 0.0 , and an identity equal or higher than 30% (Pearson, 2013). Mauve under ProgressiveMauve (default setting) (Darling et al., 2004; Darling et al., 2010) was used to compare syntenic regions. MobileElementFinder (V1.0.3) (<https://cge.cbs.dtu.dk/services/MobileElementFinder/>) was performed to predict mobile genetic elements (MGEs) associated with antibiotic and virulence factors (Minimum alignment coverage = 95; minimum sequence identity-90%, maximum truncation-30 bp) (Liu et al., 2019; Johansson et al., 2021).

Analysis of Molecular Evolution

Genetic Algorithm Recombination Detection (GARD) was used to detect recombinations. Random effect likelihood (FEL), Fast Unconstrained Bayesian Approximation (FUBAR),

and Single-likelihood ancestor counting (SLAC) methods implemented by HyPhy 2.5 under MG94xREV codon model were used to identify sites under positive selection (Guindon et al., 2010; Kosakovsky Pond et al., 2006; Kosakovsky Pond et al., 2005; Murrell et al., 2013). Evidence of episodic positive selection was verified by the Mixed-effects model of evolution (MEME) and Branch-site unrestricted statistical test for episodic diversification (BUSTED) (Kosakovsky Pond et al., 2020; Murrell et al., 2012; Murrell et al., 2015). The score values were significant whether it is less than or equal to 0.05 (BUSTED, SLAC, FEL, and MEME models), and for posterior probability if it is $\geq 95\%$ (FUBAR).

3. Results and Discussions

3.1 *Pan-genome*

Pan-genome is opened and has about 4232 protein-coding genes. The core genome has 814 protein-coding genes, and the accessory genome presents 3455 genes. Most of which are related to the metabolism of carbohydrates and amino acids (KEGG). Zou et al. (2019) described the same pattern in other bacteria of gut microbiota. According to COG distribution, most of the genes of the flexible gene pool (accessory and unique) are in the ‘Information storage and processing’ category [Translation, ribosomal structure and biogenesis (J); RNA processing and modification (A); Transcription (K); Replication, recombination and repair (L); Chromatin structure and dynamics (B)]. Analysis of USEARCH results indicated that many flexible genes encode truncated proteins. These genes might have been acquired by HGT because they often are in mobile element regions. Orthologous genes distributed in their COG and KEGG general categories are in Additional files, Figure S1. COG distribution in specific categories is in Figure S2.

KEGG analysis also displayed genes related to human infections, mainly in the flexible genome. These genes often encode proteins with bacterial metabolic functions but may be related to virulence in specific bacteria. For example, molecular chaperone DNAK protein, encoded by the DNAK gene, plays a role in stress tolerance and cell growth, but it is also a virulence factor in *Mycobacterium tuberculosis* (Tomoyasu et al., 2011; Singh et al., 2007; Fourie and Wilson, 2020). Figure 1 indicates the distribution of pan-genome across KEGG categories and a table that highlights genes associated with infectious diseases.

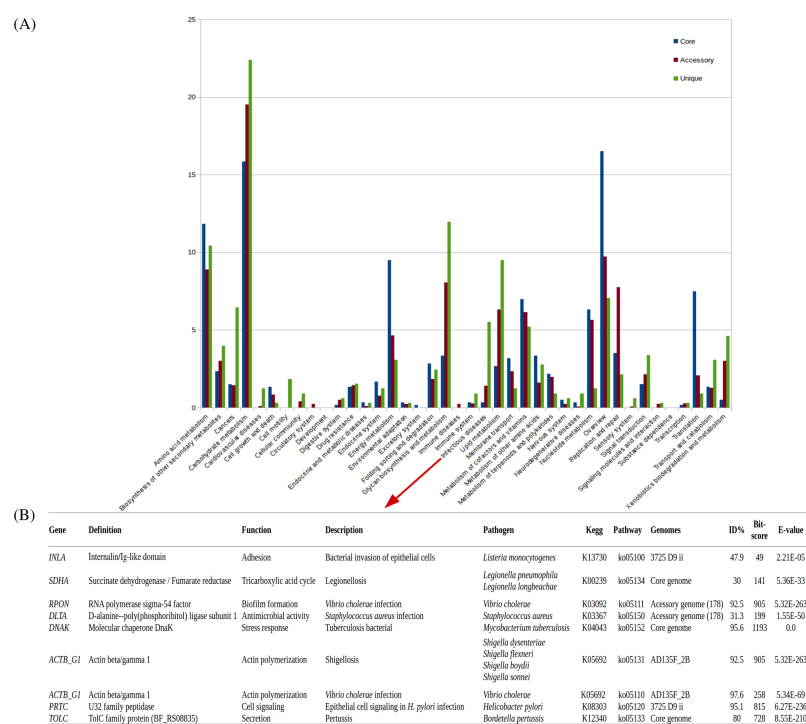


Figure1:

Distribution of orthologous genes into KEGG categories. (A) Frequency of core (in blue), accessory (in red), and unique (green) genes. The arrow points to the table corresponding to the graph bars. (B) Table showing the genes supposedly related to infectious diseases and their features.

Don PUL operon is in all strains, in this way belongs to the core genome. It encodes proteins that enable the deglycosylate N-linked glycans of the gut mucus layer, transferrin and other glycoproteins (Cao et al., 2014). Therefore, the operon is essential for *B. fragilis* adaptation to intra- and extra-intestinal environments. Other genes related to adaptation and virulence are in the accessory genome, such as *BFTP*, perforin, LPS locus, and *T6SS* genes.

Forty strains have *BFTP-1* allele while *BFTP-2* is in eight strains (20793-3, 2-078382-3, 86-5443-2-2, BOB25, CL07T00C01, CL07T12C05, CM13 and DS71). No strain has the *BFTP-3*. BfPAI is absent in 47 strains described as ETBF by Science (2013), suggesting it was not sequenced or annotated in these strains. Pan-genome prediction and the genomic location of BfPAI are in additional files, Table S2.

The perforin gene is in 80 strains, both ETBFs and NTBFs (non-enterotoxigenic *B. fragilis*). The gene encodes a nonfunctional protein in five strains (1007-1-F #8, 3783N1-2, 3783N2-1, 3976T7, and S13 L11) because of the absence of 139 amino acids from the

terminal portion of protein, which includes the MACPF domain (YGTHVLTDITLGG)-critical for protein activity (Chatzidaki-Livanis et al., 2014). Perforin [membrane attack complex/perforin (MACPF) domain] is a bacteriocin that can rupture the plasma membrane of a cell host and might be a virulence factor (Xu et al., 2010). It is found mainly in eukaryotes, and some bacteria, including *Bacteroides* spp. (Xu et al., 2010; Chatzidaki-Livanis et al., 2014).

LPS locus was described in 638R and NCTC9343 strains (Comstock et al., 1999). We also reported it in other thirteen strains [DCMOUH0042B, 3719 A10, 3986 N3, 3986 T (B) 13, 3986 T (B) 9, AD126T_1B, AD126T_2B, ATCC 25285, BFR_KZ09, CL04T03C20, HMW 615, KLE1257, and KLETY1758]. Most of them were isolated from extraintestinal infection (abscess, blood, or tissue). The capsular polysaccharide complex is one of the major virulence factors related to abscess formation (Onderdonk et al., 1977; Comstock et al., 1999), and this might explain the source of the strains.

Fifteen complete genomes have GA3 T6SS genetic architecture (Chatzidaki-Livanis et al., 2016), except the genomes of DCMOUH0067B, DCMSKEJBY0001B, FDAARGOS_763, Q1F2 strains, which belong to the carbapenem-resistant *B. fragilis* group (Valdezate et al., 2021). This finding suggests T6SS is nonessential to *B. fragilis*.

3.2 Prediction of CRISPR-Cas system

Results revealed that 76,4% of genomes have a CRISPR-Cas system. Fifty-eight strains have two CRISPR-Cas types in their genomes. Three types are only in six strains (AF26-6, BFR_KZ08, S14, S38L3, S38L5, and TM08-15). CRISPR-Cas system is absent in most NTBFs, whereas the IIIB type is the most common among ETBFs, corroborating Tajkarimi and Wexler (2017). Monophyletic groups often share the same CRISPR-Cas system. A phylogenetic tree displaying the CRISPR-Cas type of each strain is in the additional file, Figure S3.

3.3 Antimicrobial Resistance Genes

The most frequent antimicrobial resistance genes in *B. fragilis* are *CEPA* [Class A beta-lactamase (EC:3.5.2.6)] and *TETQ* (Tetracycline-resistant ribosomal protection protein). Only 40 strains presented a 'perfect match' (100%) for *CEPA*, but the results showed high

identity values in 158 strains (range from 98 to 99%, ARO: 3003559). Thus, the functionality of *CEPA* gene is probable.

Thirty-six strains have the *TETQ* gene with a ‘perfect match’, and 125 strains presented a gene with a ‘strict match’ (ARO: 3000191), but likely functional, according to comparison to protein described by Lépine et al. (1993). *TETQ* is the most abundant gene for antimicrobial in the human gut microbiota, as reported in a metagenomics analysis (Hu et al., 2013). Therefore, the spread of the gene seems advantageous.

ERMF gene [23S rRNA (Adenine (2058)-N (6))-methyltransferase (EC: 2.1.1.-)] is in 43 strains with a ‘strict match’ (identities from 98 to 99.9%), but their binding sites - essential for protein functionality (Halula et al., 1991), are unchanged. This gene confers multidrug-resistance (macrolide, lincosamide, and streptogramin B) (Park et al., 2010) and has a clinical impact on treatments. Antimicrobial resistance genes present in the strains are in additional files, Table S3.

AADS [Aminoglycoside 6-nucleotidyltransferase (EC: 2.7.1.95)], *APHA* [Aminoglycoside 3'-phosphotransferase (EC: 2.7.1.95)], *CCRA* [Subclass B1 beta-lactamase (EC 3.5.2.6)], *OXA-347* [Beta-lactamase class D OXA-209 (EC: 3.5.2.6)], *SUL2* [Dihydropteroate synthase type-2 (EC 2.5.1.15)], and *TEM-1* [Beta-lactamase class A TEM (EC: 3.5.2.6)] genes had 'perfect matches', but they were restricted to a few strains.

3.4 Mobilome

a) Analysis of Prophages

Results indicated the complete genomes have two (BOB25 strain) to fifteen (DCMOUH0067B and FDAASGOS_763 strains) prophage regions. Most virulence-related proteins in these regions had low scores, except for EF-Tu, which had higher values [VFDB database = VFG046475 (F7308_0636) Tu translation elongation factor (EF-Tu) (*Francisella* sp. TX077308) ID = 64%; bit-score = 540, and E-value = e-154]. EF-Tu protein performs various functions in bacteria, including to catalyze the binding of aminoacyl-tRNA to the ribosome, to form biofilm, and to decrease the host immune response (Harvey et al., 2019). It is unclear whether it plays a role related to virulence in *B.*

fragilis. We found no antimicrobial resistance genes in the prophage regions. The results are in the additional files, Table S4.

b) ICEFinder Analysis

The size of ICEs varies from 24.6 kilobase pairs (kb) to 193 kb, while EMIs are smaller (5 kb to 76.8 kb). Genes for antimicrobial resistance, mainly *TETQ*, were identified in ICEs and IMEs.

Most complete genomes have at least one putative ICE with T4SS, except those from BE1 and Q1F2 with only putative IMEs. Table S5, in the additional file, indicates the genomic location of the ICEs, IMEs and the antibiotic resistance genes found in them.

ICEFinder, Mauve, and NCBI BLAST search (Bit-score = 70088; e-value = 1.020e + 05; ID = 98%) analyses supported the hypothesis of an HGT (horizontal gene transfer) event between the genomic regions of *B. fragilis* DCMOUH0085B (position 2993358 to 3044318) and one from *Butyricimonas faecalis* H184 (Accession no. CP032819.1). The two regions also shared the genes for antimicrobial resistance *AADE* (aminoglycoside 6-adenylyltransferase gene), *ERMB* [rRNA adenine N-6-methyltransferase gene], (*APH* (3') - IIIA {ANT-like pseudogene (9) ANT-like aminoglycoside (9) aminoglycoside nucleotidyltransferase family gene}. HGT among Bacteroidales species was also reported by Coyne et al. (2014) and García-Bayona et al. (2021). The results of the Mauve and ICEFinder analyses are in Figure 2.

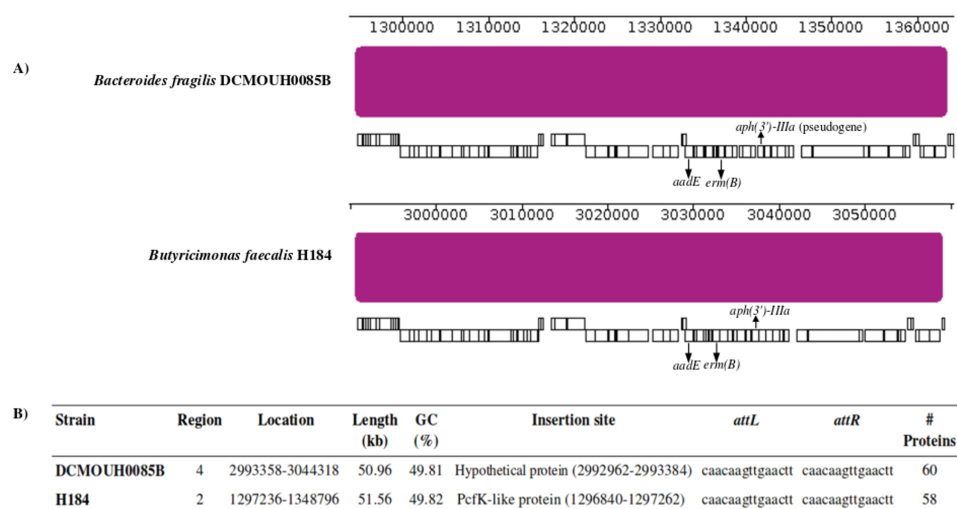


Figure 2: Comparison between two putative ICEs with T4SS. (A) Conserved synteny between their genomic regions. Antimicrobial resistance genes are indicated in each figure. (B) Features of putative ICEs with T4SS of *Bacteroides fragilis* DCMOUH0085B and *Butyrivibrio faecalis* H184.

A genomic region from the CL03T12C07 strain (position 3350878 to 3378618) is a putative ICE with T4SS that includes a CtnDOT but its *ERMF* gene is unfunctional. CtnDOT is a transposon widespread among *Bacteroides* spp. and contains the genes for antimicrobial resistance *ERMF*, *TETQ*, *TETX* (flavin-dependent monooxygenase gene) and *TETX1* (flavin-dependent monooxygenase gene) (Whittle et al., 2001; Shoemaker et al., 1989; Delavat et al., 2017).

We also identified in DCMSKEJBY0001B strain a putative ICE without identified DR that harbours a CTn341. This conjugative transposon is typical of *B. fragilis* and carries *TETQ* gene (Bacic et al., 2005). CTn341 (or a CTn341-like) may also be found in other Bacteroidates, such as *Bacteroides* sp. PHL 2737, *B. xylanisolvens* XB1A, and *Phocaeicola vulgatus* VIC01is. Thus, CTn341 seems to be the major disseminator of *TETQ* gene within order. Comparative analysis of CTn341 is in the additional file, Figure S4.

c) Conjugative Transposon

Among the conjugative transposons previously described in *Bacteroides* spp., we reported CTnDOT (aforementioned), CTn86, and CTn341.

We have identified a CTn86-like (Accession no. AY372755.1) in thirteen ETBFs strains (2-F-2#5, 3397 N3, 078320-1, 1001285H_161024_D4, AM31-13AC, BFR_KZ09, CL07T00C01, CL07T12C05, CL05T12C13, HMW 615, HMW 616, TL139C_1B, and TL139C_1B2). The *PER* gene, which encodes an ABC permease transporter protein, was excluded from these strains, possibly because of transposon insertion, as described by Franco (2004). CTn86 also was identified in 20793-3, 86-5443-2-2 and BOB25 strains (Pierce and Bernstein, 2016; Franco, 2004; Nikitina et al., 2015).

CTn341-like is in 20656-2-1, AF14-14AC and AF14-26 strains. It does not have all CTn341 genes and attachment sites (*attL* and *attR*) (Guédon et al., 2017). It has genes for transfer (*TRAG*, *TRAI*, *TRAJ*, and *TRAM* genes), integration (*INT* gene), mobilization (*MOBA*, *MOBB*, and *MOBC*) and excision (*EXC* gene) (Bonheyo et al., 2001; García et al.,

2009; Haggoud et al., 1994), besides *TETQ*. Therefore, its functionality is uncertain. Comparative analysis of CTn341 is in the additional file, Figure S5.

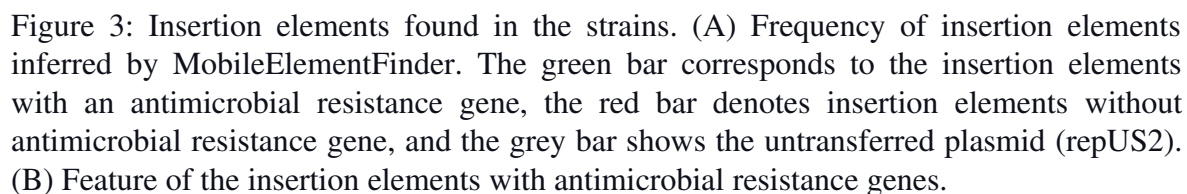
c) Insertion Elements (IS)

ISBf5 (IS1182 family) is the insertion element most common in *B. fragilis* genomes. It has 1830 bp and carries one transposase gene (Kuwahara et al., 2004).

ISBf1 (IS21 family) (Rogers et al., 1994) is in some strains, but its *CEPA* gene is incomplete in eight strains (885_BFR, AD135F_2B, AF32-10, Am47-7, COR2-248-WT-1, HCK-B3, HAP130N_2B, TM08 -15) and absent in AD135F_1B, AD135F_3B, ATCC 25285, and Ds-233 strains. However, they have in their genomes a *CEPA* functional with insertion of IS1224 in their promoter region that increases *CEPA* expression (Rogers et al., 1994; Vandecraen et al., 2017).

Functional resistance genes were identified in ISBthe1 (*NIMB* gene, metronidazole), ISBf8 (*CFXA* gene), and IS4351 (*TETX* gene). ISBf8 and IS4351 are within the non-conjugative transposons (Tn4555 and Tn4351, respectively), and for that reason, they cannot spread their genes.

Therefore, the spread of genes for antibiotic resistance in *B. fragilis* depends on conjugative transposons, as suggested by Whittle et al. (2002) and Johnson and Grossman (2015), and also of ICE with T4SS, as we observed in our results. IS frequency and its genes for antimicrobial resistance are in Figure 3. The genomic location of the insertion elements is in the additional file, Table S6.



a) Molecular Evolution of Antibiotic Resistance Genes

TETQ alignment displayed 116 variable sites. FUBAR method inferred three codon sites are under positive selection: 35 (posterior probability = 0.9676), 237 (posterior probability = 0.9669), and 426 (posterior probability = 0.9623) while MEME suggested episodic positive selection at 354 codon (p-value = 0.003), and 550 (p-value = 0.0464). Comparative analysis with the TETQ protein (described by Lépine et al. 1993) indicates that the four GDP-GTP binding sites are intact.

b) Molecular Evolution in Pan-genome

At least 396 genes are under positive selection (episodic or pervasive). Most of them encode membrane proteins, such as adenosine triphosphate (ATP)-binding cassette (ABC) superfamily (*MACB* genes), resistance-nodulation-division (RND) family (*MACA*, *MDTA*, *MDTE*, and *MEXA* genes), and TonB-dependent receptor (*SUSC* and *SUSD* genes).

ATP-binding and RND proteins are efflux pumps that facilitate the elimination of toxins and are related to antimicrobial resistance phenotypes (Li and Nikaido, 2004; Pimenta et al., 2005; Piddock, 2006; Davidson et al., 2008).

TonB-dependent receptors play a role in transport and nutrient absorption (Hashimoto et al., 2005; Yoshizaki et al., 2014). Recently, they have been the subject of research for the development of vaccines against Gram-negative bacteria because of their cell multiple functions (Wang et al., 2021).

Seven paralogous genes of the Colicin I receptor (*CIRA* genes) also had several sites under positive selection. These proteins belong to the TonB-dependent copper receptor family (IPR010100) and perform in the transport of nutrients across the periplasmic space (Chimento et al., 2003; Reeves et al., 1996). Colicins in *Escherichia coli* are toxins that act in intraspecific competition (Cascales et al., 2007), but their role in *B. fragilis* is unclear. Yoshizaki et al. (2014) compared the genomes of 638R, NCTC9343, and YCH46, and found 52 genes under positive selection. The greater diversity used in our work might explain the difference between the results. The results of Hyphy are in the additional file, Table S7.

4. Conclusions

HGT is an event of extreme evolutionary relevance for *B. fragilis*. *TETQ* gene spread was mainly by ICE with T4SS and conjugative transposons, whereas *the CEPA* gene was an ancient spread through a mobile element that lost its function. We can describe HGT events not yet reported in previous research. Nonetheless, experimental validation is required. Besides the antimicrobial resistance genes, other genes might guarantee adaptive advantage for *B. fragilis*. They are related to efflux pumps facilitating the exit of substances from the cells (*MACB*, *MACA*, *MDTA*, *MDTE*, and *MEXA* genes); transport and absorption of nutrients (*SUSC* and *SUSD* genes); and alternative carbon acquisition (DON

PUL). There is no evidence that *EF-Tu*, *DNAK* or colicin genes are associated with virulence and need further research to clarify their functions in *B. fragilis*.

Abbreviation:

AMR: Antimicrobial multidrug-resistance

BFT: Fragilysin

CRISPR: Clustered regularly interspaced short palindromic repeat

ETBF: Enterotoxigenic *Bacteroides fragilis*

HGT: Horizontal gene transfer

ICE: Integrative and conjugative elements

IME: Integrative mobilizable elements

NTBF: Non-enterotoxigenic *B. fragilis*

Acknowledgments

This manuscript is part of the Ph.D. thesis of Eliane Evanovich from Programa de Pós-graduação em Genética e Biologia Molecular, Universidade Federal do Pará. The research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors wish to express their thanks for the financial support of Pró-Reitoria de Pesquisa e Pós-Graduação (PROPESP).

Conflicts of Interest

The authors declare that they have no competing interests.

Authors Contribution

EE conducted the analysis, designed and wrote the paper. PJSMM revised the document and JFG supervised the study and fully revised the document. All authors read and approved the final version.

References

Adegbola RA (1988) Bacterial adhesion and pathogenicity. Afr J Med Med Sci 17:63-69.

Akhter S, Aziz RK, and Edwards RA (2012) PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. *Nucleic Acids Res* 40:e126.

Alcock BP, Raphenya AR, Lau TTY and Tsang KK (2020) CARD 2020: Antibiotic Resistome Surveillance with the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res* 48:D517-D525.

Altschul, S, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, and Lipman, DJ (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.

Anka N, Sydenhamb TV, Iversenc LH, Justesen US, Wang M (2015) Characterisation of a multidrug-resistant *Bacteroides fragilis* isolate recovered from blood of a patient in Denmark using whole-genome sequencing. *Int J Antimicrob Agents* 46:117-120.

Bacic M, Parker AC, Stagg J, Whitley HP, Wells WG, Jacob LA, Smith CJ (2005) Genetic and structural analysis of the *Bacteroides* conjugative transposon CTn341. *J Bacteriol* 187:2858-69.

Bamias G and Cominelli F (2016) Cytokines and intestinal inflammation. *Curr Opin Gastroenterol* 32:437-442.

Bonheyo GT, Hund BD, Shoemaker NB and Salyers AA (2001) Transfer region of a *Bacteroides* conjugative transposon contains regulatory as well as structural genes. *Plasmid* 46:202–209.

Cao Y, Rocha ER and Smith CJ (2014) Efficient utilization of complex N-linked glycans is a selective advantage for *Bacteroides fragilis* in extraintestinal infections. *Proc Natl Acad Sci U S A* 111:12901–12906.

Cascales E, Buchanan SK, Duché D, Kleanthous C, Lloubès R, Postle K, Riley M, Slatin S and Cavard D (2007) Colicin biology. *Microbiol Mol Biol Rev* 71:158–229.

Chatzidaki-Livanis M, Coyne MJ, Comstock LE (2014) An antimicrobial protein of the gut symbiont *Bacteroides fragilis* with a MACPF domain of host immune proteins. *Mol Microbiol* 94:1361–1374.

Chatzidaki-Livanis M, Geva-Zatorsky N, Comstock LE (2016) *Bacteroides fragilis* type VI secretion systems use novel effector and immunity proteins to antagonize human gut Bacteroidales species. *Proc Natl Acad Sci U S A* 113:3627–3632.

Chaudhari NM, Gupta VK and Dutta C (2016) BPGA- an ultra-fast pan-genome analysis pipeline. *Sci Rep* 6:24373.

Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, and Jin Q (2005) VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res* 33:D325-8.

Cheng WT, Kantilal HK, Davamani F (2020) The Mechanism of *Bacteroides fragilis* toxin Contributes to Colon Cancer Formation. *Malays J Med Sci* 27:9–21.

Chimento DP, Mohanty AK, Kadner RJ and Wiener MC (2003) Substrate-induced transmembrane signaling in the cobalamin transporter Btu. *Nat Struct Biol* 10:394-401.

Comstock LE, Coyne MJ, Tzianabos AO and Kasper DL (1999) Interstrain Variation of the Polysaccharide B Biosynthesis Locus of *Bacteroides fragilis*: Characterization of the Region from Strain 638R. *J Bacteriol* 181:6192–6196.

Couvin D et al. (2018) CRISPRCasFinder, an update of CRISPRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res* 46:W246–W251.

Coyne MJ, Tzianabos AO, Mallory BC, Carey VJ, Kasper DL and Comstock LE (2001) Polysaccharide Biosynthesis Locus Required for Virulence of *Bacteroides fragilis*. *Infect Immun* 69:4342–4350.

Coyne MJ, Zitomersky NL, McGuire AM, Earl AM and Comstock LE (2014). Evidence of Extensive DNA Transfer between Bacteroidales Species within the Human Gut. *mBio* 5: e01305-14.

Darling AE, Mau B, and Perna NT (2010) ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5: e11147.

Darling AE, Mau B, Blattner FR, and Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403.

Davidson AL, Dassa E, Orelle C and Chen J (2008) Structure, function, and evolution of bacterial ATP-binding cassette systems. Microbiol Mol Biol Rev 72:317-364.

Delavat F, Miyazaki R, Carraro N, Pradervand N, van der Meer R (2017) The hidden life of integrative and conjugative elements. FEMS Microbiology Reviews 41: 512–537.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797.

Elsaghir H and Reddivari AKR (2021) *Bacteroides fragilis*, <https://www.ncbi.nlm.nih.gov/books/NBK553032/> (March 18, 2021)

Finegold SM (1995) Overview of clinically important anaerobes. Clin Microbiol Rev 20:205-207.

Fourie KR and Wilson HL (2020) Understanding GroEL and DnaK Stress Response Proteins as Antigens for Bacterial Diseases. Vaccines 8:773.

Franco AA (2004) The *Bacteroides fragilis* pathogenicity island is contained in a putative novel conjugative transposon. J Bacteriol 186:6077-6092.

Franco AA, Cheng RK, Goodman A, Sears CL (2002) Modulation of bft expression by the *Bacteroides fragilis* pathogenicity island and its flanking region. Mol Microbiol 45:1067-77.

Franco AA, Mundy, LM, Trucksis M, Wu S, Kaper JB and Sears CL (1997) Cloning and characterization of the *Bacteroides fragilis* metalloprotease toxin gene. Infect. Immun 65:1007-1013.

García N, Gutiérrez G, Lorenzo M, Vadillo S, Píriz S and Quesada A (2009) Gene Context and DNA rearrangements in the carbapenemase locus of division II strains of *Bacteroides fragilis*. *Antimicrob Agents Chemother* 53:2677-8.

García-Bayona L, Coyne MJ and Comstock LE (2021) Mobile Type VI secretion system loci of the gut Bacteroidales display extensive intra-ecosystem transfer, multi-species spread and geographical clustering. *PLoS Genet* 17:e1009541.

Guédon G, Libante V, Coluzzi C, Payot S and Leblond-Bourget N (2017) The Obscure World of Integrative and Mobilizable Elements, Highly Widespread Elements that Pirate Bacterial Conjugative Systems. *Genes (Basel)* 8:337.

Guglielmini J, Quintais L, Garcillán-Barcia MP, de la Cruz F, Rocha EP (2011) The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. *PLoS Genet* 7:e1002222.

Guindon S, Dufayard JF and Lefort V, Anisimova M, Hordijk W and Gascuel O (2010) New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst Biol* 59:307-321.

Haggoud A, Reysset G, Azeddoug H and Sebald M (1994) Nucleotide sequence analysis of two 5-nitroimidazole resistance determinants from *Bacteroides* strains and of a new insertion sequence upstream of the two genes. *Antimicrob Agents Chemother* 38:1047-51.

Halula MC, Manning S and Macrina FL (1991) Nucleotide sequence of *ermFU*, a macrolide-lincosamide-streptogramin (MLS) resistance gene encoding an RNA methylase from the conjugal element of *Bacteroides fragilis* V503. *Nucleic Acids Res* 19:3453.

Harvey KL, Jarocki VM, Charles IG and Djordjevic SP (2019) The Diverse Functional Roles of Elongation Factor Tu (EF-Tu) in Microbial Pathogenesis. *Front Microbiol* 10:2351.

Herrou J, Choi VM, Wardenburg JM and Crosson S (2016) Activation Mechanism of the *Bacteroides fragilis* Cysteine Peptidase, Fragipain. *Biochemistry* 55:4077–4084.

Ho BT, Dong TG and Mekalanos JJ (2014) A view to a kill: the bacterial type VI secretion system. *Cell Host Microbe* 15:9–21.

Hu Y, Yang X, Qin J, Lu N, Cheng G, Wu N, Pan Y, Li J, Zhu L, Wang X, et al. (2013) Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat Commun* 4:2151.

Jia B, Raphenya A, Alcock B, Wagglechner N, Guo P, Tsang K, Lago B, Dave B, et al. (2017) CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res* 45:D566-D573.

Johansson MHK, Bortolaia V, Tansirichaiya S, Aarestrup FM, Roberts AP, and Petersen TN (2021) Detection of mobile genetic elements associated with antibiotic resistance in *Salmonella enterica* using a newly developed web tool: MobileElementFinder. *J Antimicrob Chemother* 76:101–109.

Johnson CM, Grossman AD (2015) Integrative and Conjugative Elements (ICEs): What They Do and How They Work. *Annu Rev Genet* 49: 577-601.

Kato N, Liu CX, Kato H, Watanabe K, Tanaka Y, Yamamoto T, Suzuki K and Ueno K (2000) A new subtype of the metalloprotease toxin gene and the incidence of the three bft subtypes among *Bacteroides fragilis* isolates in Japan. *FEMS Microbiol Lett* 182:171-176.

Ko SH, Rho DJ, Jeon JI, Kim Y, Woo HE, Lee YK and Kim JM (2016) *Bacteroides fragilis* Enterotoxin Upregulates Heme Oxygenase-1 in Intestinal Epithelial Cells via a Mitogen-Activated Protein Kinase- and NF-kappaB-Dependent Pathway, Leading to Modulation of Apoptosis. *Infect Immun* 84:2541–2554.

Kosakovsky Pond SL and Frost SDW (2005) Not So Different After All: A Comparison of Methods for Detecting Amino Acid Sites Under Selection. *Mol Biol Evol* 22:1208-1222.

Kosakovsky Pond SL, Poon AFY, Velazquez R, Weaver S, Hepler NL, Murrell B, Shank DL, Magalis BR, Bouvier D, Nekrutenko A(2020) HyPhy 2.5—A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Mol Biol Evol* 37:295-299.

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH and Frost SD (2006) Automated Phylogenetic Detection of Recombination Using a Genetic Algorithm. *Mol Biol Evol* 23:1891-1901.

Kuwahara T, Yamashita A, Hirakawa H, Nakayama H, Toh H, Okada N, Kuhara S, Hattori M, Hayashi T. and Ohnishi Y. (2004) Genomic analysis of *Bacteroides fragilis* reveals extensive DNA inversions regulating cell surface adaptation. *PNAS* 101: 14919-14924

Lépine G, Lacroix JM and Walker CB (1993) A Progulske-Fox. Sequencing of a *tet(Q)* gene isolated from *Bacteroides fragilis* 1126. *Antimicrob Agents Chemother* 37:2037-2041.

Li XZ and Nikaido H (2004) Efflux-mediated drug resistance in bacteria. *Drugs* 64:159-204.

Liu M, Li X, Xie Y, Bi D, Sun J, Li J, Tai C, Deng Z, and Ou H (2019) ICEberg 2.0: an updated database of bacterial integrative and conjugative elements. *Nucleic Acids Res* 47:D660–D665.

Mazmanian SK, Round JL and D.L Kasper (2008) A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* 453:620–625.

Moncrief JS, Duncan AJ, Wright RL, Barroso LA and Wilkins TD (1998) Molecular characterization of the fragilysin pathogenicity islet of enterotoxigenic *Bacteroides fragilis*. *Infect Immun* 66:1735-1739.

Murrell B, Weaver S, Smith MD, Wertheim JO, Murrell S, Aylward A, Eren K, Pollner T, Martin DP, Smith DM, et al. (2015) Gene-Wide Identification of Episodic Selection. *Mol Biol Evol* 32:1365-1371.

Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond S and Scheffler K (2013) FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol* 30:1196-1205.

Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K and Kosakovsky Pond SL (2012) Detecting Individual Sites Subject to Episodic Diversifying Selection. *PLoS Genetics* 8:e1002764.

Niestępski S, Harnisz M, Korzeniewska E, Osińska A and Dziuba B (2017) *Bacteroides spp.* - clinical significance, antibiotic resistance and identification methods. *Postepy Mikrobiol* 56:67–76.

Nikitina AS, et al. (2015) Complete Genome Sequence of an Enterotoxigenic *Bacteroides fragilis* Clinical Isolate. *Genome Announc* 3:e00450-15.

Onderdonk AB, Kasper DL, Cisneros RL and Bartlett JG (1977) The capsular polysaccharide of *Bacteroides fragilis* as a virulence factor: comparison of the pathogenic potential of encapsulated and unencapsulated strains. *J Infect Dis* 136:82-89.

Papaparaskevas J, Katsandri A, Pantazatou A, Stefanou I, Avlamis A, Legakis NJ and Tsakris A (2011) Epidemiological characteristics of infections caused by *Bacteroides*, *Prevotella* and *Fusobacterium* species: A prospective observational study. *Anaerobe* 17:113-117.

Park EK, Kim H and Jin HJ (2010) Phylogenetic analysis of rRNA methyltransferases, Erm and KsgA, as related to antibiotic resistance. *FEMS Microbiol Lett* 309:151–162.

Pearson WR (2013) An introduction to sequence similarity ("homology") searching. *Curr Protoc Bioinform* 3:3.1.

Piddock LJ (2006) Multidrug-resistance efflux pumps - not just for resistance. *Nat Rev Microbiol* 4:629-36.

Pierce JV and Bernstein HD (2016) Genomic Diversity of Enterotoxigenic Strains of *Bacteroides fragilis*. *PLoS One* 11:e0158171.

Pimenta AL, Racher K, Jamieson L, Blight MA and Holland IB (2005) Mutations in HlyD, part of the type 1 translocator for hemolysin secretion, affect the folding of the secreted toxin. *J. Bacteriol* 187:7471-80.

Reeves AR, D'Elia JN, Frias J, Salyers AA (1996) A *Bacteroides thetaiotaomicron* outer membrane protein that is essential for utilization of maltooligosaccharides and starch. *J. Bacteriol* 178:823-30.

Rogers MB, Bennett TK, Payne CM and Smith CJ (1994) Insertional activation of *CEPA* leads to high-level beta-lactamase expression in *Bacteroides fragilis* clinical isolates. J Bacteriol 176: 4376–4384.

Rogers MB, Parker AC, Smith CJ (1993) Cloning and characterization of the endogenous cephalosporinase gene, *CEPA*, from *Bacteroides fragilis* reveals a new subgroup of Ambler class A beta-lactamases. Antimicrob Agents Chemother 37: 2391–400.

Round JL and Mazmanian SK (2010) Inducible Foxp3p regulatory T-cell development by a commensal bacterium of the intestinal microbiota. Proc Natl Acad Sci U S A 107:12204–12209.

Russell AB, Hood RD, Bui NK, LeRoux M, Vollmer W and Mougours JD (2011) Type VI secretion delivers bacteriolytic effectors to target cells. Nature 475:343–347.

Science, U.o.M.S.o.M.I.f.G. (2013). GSCID Genomic Sequencing Center for Infectious Disease [Online]. NCBI. Available online at: http://gscid.igs.umaryland.edu/wp.php?wp=comparative_genomics_of_human_mucosal_isolates_of_the_enterotoxigenic_bacteroides_fragilis_group.

Sears CL (2001) The toxins of *Bacteroides fragilis*. Toxicon 39:1737–1746.

Sears CL, Myers LL, Lazenby and Van Tassell RL (1995) Enterotoxigenic *Bacteroides fragilis*: a Rogue among Symbiotes. Clin Microbiol Rev 2:349–69.

Shiryaev SA, Aleshin AE and Muranaka N (2014) Structural and functional diversity of metalloproteinases encoded by the *Bacteroides fragilis* pathogenicity island. FEBS J 281:2487–2502.

Shoemaker NB, Barber RD and Salyers AA (1989) Cloning and characterization of a *Bacteroides* conjugal tetracycline-erythromycin resistance element by using a shuttle cosmid vector. J Bacteriol 171:1294–1302.

Sims D, Sudbery I, Ilott NE, Heger A and Ponting CP (2014) Sequencing depth and coverage: key considerations in genomic analyses. Nat Rev Genet 15:121–132.

Singh VK, Utaida S, Jackson LS, Jayaswal RK, Wilkinson BJ and Chamberlain NR (2007) Role for *dnaK* locus in tolerance of multiple stresses in *Staphylococcus aureus*. *Microbiology* 153:3162–3173.

Tajkarimi M and Wexler HM (2017) CRISPR-Cas Systems in *Bacteroides fragilis*, an Important Pathobiont in the Human Gut Microbiome. *Front Microbiol* 23:2234, 2017.

Tomoyasu T, Tabata A, Imaki H, Tsuruno K, Miyazaki A, Sonomoto K, Whiley RA and Nagamune H (2011) Role of *Streptococcus intermedius* DnaK chaperone system in stress tolerance and pathogenicity. *Cell Stress Chaperon* 17:41–55.

Vacante M, Ciuni R, Basile F and Biondi A (2020) Gut Microbiota and Colorectal Cancer Development: A Closer Look to the Adenoma-Carcinoma Sequence. *Biomedicines* 8:489.

Valdezate S, Cobo F, Monzón S, Medina-Pascual MJ, Zaballos Á, Cuesta I, Pino-Rosa S and Villalón P (2021) Genomic Background and Phylogeny of *cfiA*-Positive *Bacteroides fragilis* Strains Resistant to Meropenem-EDTA. *Antibiotics (Basel)* 10:304.

Valguarnera E and Wardenburg JB (2020) Good gone bad: One toxin away from disease for *Bacteroides fragilis*. *J Mol Biol* 432:765-785.

Vandecraen J, Chandler M, Aertsen A and Van Houdt R (2017) The impact of insertion sequences on bacterial genome plasticity and adaptability. *Crit Rev Microbiol* 43:709-730.

Vestby LK, Grønseth T, Simm R and Nesse LL (2020) Bacterial Biofilm and its Role in the Pathogenesis of Disease. *Antibiotics* 2:59.

Wang J, Xiong K, Pan Q, He W, Cong Y (2021). Application of TonB-Dependent Transporters in Vaccine Development of Gram-Negative Bacteria. *Front Cell Infect Microbiol* 10:589115.

Wexler HW (2007) *Bacteroides*: the Good, the Bad, and the Nitty-Gritty. *Clin Microbiol Rev* 20:593-621.

Whittle G, Hund BD, Shoemaker NB, Salyers AA (2001) Characterization of the 13-kilobase *ermF* region of the *Bacteroides* conjugative transposon CTnDOT. *Appl Environ Microbiol* 67:3488-95.

Whittle G, Shoemaker NB and Salyers AA (2002) The role of *Bacteroides* conjugative transposons in the dissemination of antibiotic resistance genes. *Cell Mol Life Sci* 59:2044-54.

Williams T and Kelley C (2018) Gnuplot 4.6.6, <http://www.gnuplot.info> (March, 18, 2021).

Wu S, Lim KC, Huang J, Saidi RF and Sears CL (1998) *Bacteroides fragilis* enterotoxin cleaves the zonula adherens protein, E-cadherin, *Proc Natl Acad Sci U S A* 95:14979-14984.

Xu Q, Abdubek P, Astakhova T, Axelrod HL, Bakolitsa C, Cai X, Carlton D, Chen C, Chiu HJ, Clayton T, et al. (2010) Structure of a membrane-attack complex/perforin (MACPF) family protein from the human gut symbiont *Bacteroides thetaiotaomicron*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 66:1297-305.

Yoshizaki S, Umemura T, Tanaka K, Watanabe K, Hayashi M, Muto Y (2014) Genome-wide evidence of positive selection in Bacteroides fragilis. Computational Biology and Chemistry 52: 43-50.

Zou Y, Xue W, Luo, G, Deng Z, Qin P, Guo R, Sun H, Xia Y, Liang S, Dai Y, Wan D. et al. (2019) 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nat Biotechnol* 37:179–185.

3. CAPÍTULO II

Genômica Comparativa de *Lactiplantibacillus plantarum*

Artigo publicado no periódico: *International Journal of Genomics* (Int J Genomics. 2019

Apr 10;2019:4973214. doi: 10.1155/2019/4973214. PMID: 31093491)

Research Article

Comparative Genomic Analysis of *Lactobacillus plantarum*: An Overview

Eliane Evanovich , Patricia Jeanne de Souza Mendonça Mattos, and João Farias Guerreiro

Laboratório de Genética Humana e Médica, Instituto de Ciências Biológicas-Universidade Federal do Pará, Pará, Brazil

Correspondence should be addressed to Eliane Evanovich; lianevanovich@gmail.com

Received 22 November 2018; Revised 19 February 2019; Accepted 12 March 2019; Published 10 April 2019

Academic Editor: Antonio Ferrante

Copyright © 2019 Eliane Evanovich et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. *Lactobacillus plantarum* is widely used in the manufacture of dairy products, fermented foods, and bacteriocins. The genomes of the strains contain multiple genes which may have been acquired by horizontal gene transfer. Many of these genes are important for the regulation, metabolism, and transport of various sugars; however, other genes may carry and spread virulence and antibiotic resistance determinants. In this way, monitoring these genomes is essential to the manufacture of food. In this study, we aim to provide an overview of the genomic properties of *L. plantarum* based on approaches of comparative genomics. **Results.** The finding of the current study indicates that the core genome of *L. plantarum* presents 1425 protein-coding genes and is mostly related to the metabolic process. The accessory genome has on average 1320 genes that encodes protein involved in processes as the formation of bacteriocins, degradation of halogen, arsenic detoxification, and nisin resistance. Most of the strains show an ancestral synteny, similar to the one described in the genomes of *L. pentosus* KCA1 and *L. plantarum* WCFS1. The lifestyle island analyses did not show a pattern of arrangement or gene content according to habitat. **Conclusions.** Our results suggest that there is a high rate of transfer of genetic material between the strains. We did not identify any virulence factors and antibiotic resistance genes on the genomes. Thus, the strains may be useful for the biotechnology, bioremediation, and production of bacteriocins. The potential applications are, however, restricted to particular strains.

1. Background

Lactobacillus plantarum is a facultative heterofermentative lactic acid bacteria (LAB). It may obtain energy from different sugars, and it occupies an adaptive nutrient-rich niche such as gastrointestinal, vaginal, and urogenital tracts, vegetables, dairy products, and fermented foods [1–6]. This bacteria also carries relevant properties for not only the manufacture of a variety of food and wine but also vitamin production, bacteriocin, probiotic, antifungal, and potential anticaries agents [7–11]. In general, LAB has reduced genomes, but *L. plantarum* presents a larger genome with numerous genes, which have been acquired by horizontal gene transfer (HGT) mainly via mobile elements (prophages, plasmids, transposons, and integrons) [12, 13].

L. plantarum strains are capable of producing different antimicrobial compounds, such as hydrogen peroxide, organic acids (primarily lactic and acetic acid), antiaflatoxinogenic,

and bacteriocins [14–16]. The latter act against a wide range of bacterial pathogens, in the broad and narrow spectra [17, 18]. The plantaricins (or two-peptide bacteriocins) are usually produced by *L. plantarum* and include the plantaricins J51, JK, and EF [19–21].

Kleerebezem et al. [12] and Molenaar et al. [22] described in the *L. plantarum* WCFS1 genome a region known as the lifestyle island. It may have been acquired by HGT and is divided into two subregions of approximately 150 kb and 190 kb and contains several genes critical for the regulation, metabolism, and transport of sugars.

Mobile genetic elements are segments of DNA that can move within and between bacteria. They are potential disseminators of virulence factors and determinants of antibiotic resistance (AR). Several of these elements are found in the genomes of *L. plantarum* [23–27]; therefore, it is essential to screen these unwanted genes in the new strains. Previous studies have reported that AR genes have already been

described in *L. crispatus*, *L. gasseri*, *L. reuteri*, and *L. plantarum* although their strains are considered safe in the United States through the Generally Recognized as Safe (GRAS) designation [28–35].

The aim of the paper is to provide an overview of the structural and genomic properties of *L. plantarum* genome strains available in the GenBank sequence database. We used complete genomes for avoiding the underestimation of gene content [36, 37]. The results show that the great majority of the mobile elements are Sha1 and Phig1 bacteriophages, originally isolated from *L. plantarum* [38, 39]. It suggests a high gene transfer rate between the strains. The outcomes also suggest a great potentiality in producing bacteriocins, except for the strains 16 and Zhang-LL. Therefore, the various applications of strains are unequivocal. In addition to their recognized applications, the strains may be useful to the pharmaceutical industry, in the bioremediation of halogenated pollutants and arsenic-rich soils.

2. Methods

2.1. Complete Genomes. The genome data was available in the NCBI (National Center for Biotechnology Information) (<https://www.ncbi.nlm.nih.gov>). Details regarding the identification and source of the samples used are in the additional file Table S1. Prokka version 1.12-beta [40] with the arguments kingdom Bacteria, and genus *Lactobacillus* was used for verifying the genome annotations.

2.2. Prediction of CRISPR Sequences and Mobile Elements. CRISPR sequences, mobile elements, genomic islands, and resistance genes were obtained via CRISPRFinder (<http://crispr.i2bc.paris-saclay.fr/Server/>) and PHAge Search Tool Enhanced Release (PHASTER) (<http://phaster.ca>) [40, 41] under the references intact score > 90, questionable score 70–90, and incomplete score < 70.

CARD (Comprehensive Antibiotic Resistance Database) (<http://arpcard.mcmaster.ca/>) was used to predict the resistance under the BLAST (Basic Local Alignment Search Tool), and RGI (Resistance Gene Identifier) (under Perfect hit, rigorous hit alone, and Perfect, Strict, and Loose hit criteria) [42]. The ResFinder 3.0 server (<https://cge.cbs.dtu.dk/services/ResFinder/>) was used to identify acquired antimicrobial resistance genes and/or chromosomal mutations [43].

2.3. Pan-genome. BPGA (Bacterial Pan Genome Analysis) tool [44] version 1.3 was used to identify core, accessory, and unique protein families. It was also used to search for the presence or absence of genes, phylogenetic inference, and atypical GC content and for mapping gene functions based on COG (Clusters of Orthologous Groups of proteins). The orthologous clusters were generated via USEARCH 9.2.64 (identity cut off = 50%) [45]. MUSCLE generated the alignments and the phylogenies [46], and gnuplot 4.6.6 (<https://sourceforge.net/projects/gnuplot/files/latest/download> source code freely distributed) was applied to plot the graphs [47].

2.4. Multiple Genome Alignment. Mauve, under progressive-Mauve, was used to perform the synteny analyses and the

multiple genome alignments (default setting) [48, 49]. *L. pentosus* KCA1 was used as the outgroup. Another analysis using the last subregion of the lifestyle island was also performed. In this, the value of minimal LCBs was equal to 1000.

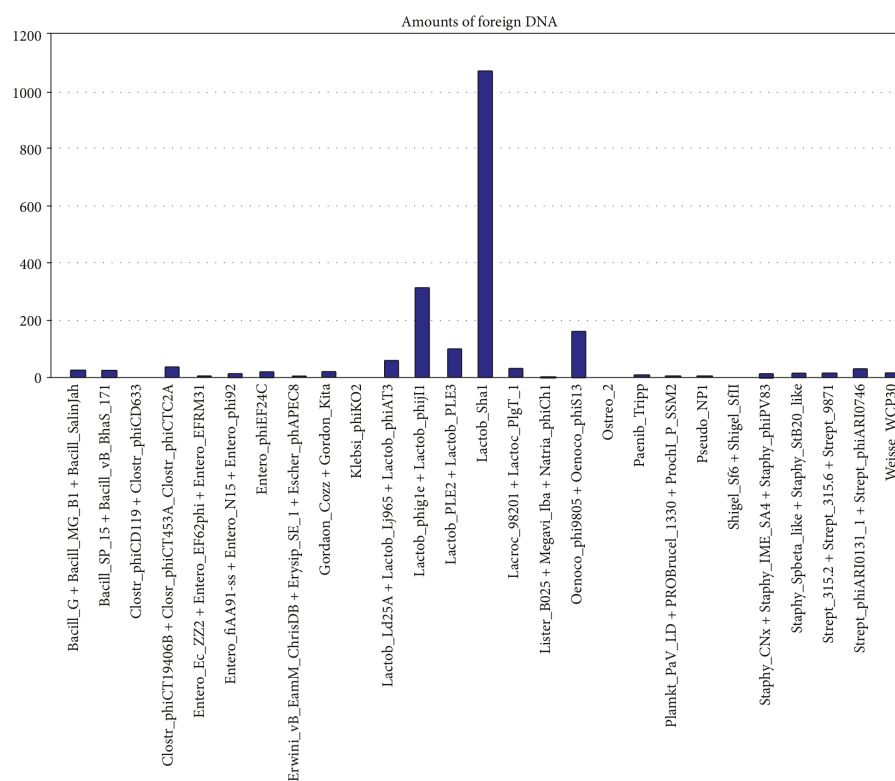
3. Results and Discussion

3.1. Mobilome and Resistome. The results obtained via PHASTER showed that the sequences of bacteriophage origin have about 151 kb, i.e., about 48% of the size of the *L. plantarum* genomes. Bacteriophage proteins (DNA packaging protein, holin protein, lysin, tail, capsid, protease, terminase, and integrase) and hypothetical proteins were the most frequent (about 91%). The bacteriophages most encountered were Sha1 and Phig1, both isolated from *L. plantarum* [34, 35] as shown in Figure 1 and the additional file Table S2.

Nine of the 49 genomes display the CRISPR-Cas system (class 2, type II with four genes, cas9, cas1, cas2, and cns2, as found in *Streptococcus thermophilus*) [50]. These strains are from fermented foods (LY-78, MF1298, ZS2058, and TS12), raw milk (LZ206 and LZ227), an environmental sample (CLP0611), faeces of a newborn (ZJ316), and a cell culture (CGMCC 1.557). Length of the CRISPR sequence varies from 300 to 2111 bp, and the number of CRISPR spacers was four to 31. The degenerate repeat DR-consensus (5'-GTCCTGAATAGTAGTCATATCAACA GGTTTAGAAC-3') was equally reported in the *L. pentosus* MP-10 and *L. pentosus* KCA1 genomes [51, 52]. Evaluation of the spacer sequences revealed several invasion events by *Lactobacillus* bacteriophages (from *L. alimentarius* DSM 20249, *L. brevis* 925A, and *L. helveticus* FAM8627) and mainly by *L. plantarum* bacteriophages. This result is consistent with what was obtained via PHASTER and suggests that the CRISPR-Cas system is not the primary defence against bacteriophage invasion (additional file: Table S3). According to Abriouel et al. [51], the presence of bacteriophages may provide some selective advantage to the bacterial cell, by helping in the fight against other prophage infections. The domestication of mobile genetic elements, which is useful for different bacterial processes, has been described [53–56], and it may also be applied to *Lactobacillus*, including *L. plantarum*.

ResFinder, CARD, and PATRIC did not indicate potential antibiotic resistance or virulence determinants. Most *L. plantarum* strains have putative genes annotated as antibiotic resistance genes or a virulence factor (such as the putative formate acetyltransferase 3-ybiW gene or mdxE-maltodextrin ABC-transporter protein gene), but they seem to be only spurious partial hits [40], which may exert other cellular functions. AR genes were not detected in the plasmids, as well.

3.2. Pan-genome. On average, the genomes present 2917 protein-coding genes, 1425 of which belong to the core genome. Most core orthologous groups (OGs) are related to metabolism. OG distribution in COG categories is shown in Figure 2. This result is not surprising, due to the lifestyle island [3, 4, 57]. Other important OGs from the core genome are protein-coding genes involved in the synthesis of

FIGURE 1: Types of prophage detected in *L. plantarum* genomes.

exopolysaccharides (EPS), histidine protein kinase (HPK), L-2-haloacid dehalogenase, sortase A (srtA), and fibrinogen-binding. They are essential genes to the synthesis of plantaricin and degradation of halogenated compounds and for host-bacterial interaction [58–64].

Hpk6, hpk7, and hpk11 proteins, belonging to histidine protein kinase (HPK), have a regulatory function in the synthesis of plantaricins and are therefore crucial to the strains [61]. The *deh* gene encodes an L-2-haloacid dehalogenase, an enzyme that degrades halogenated compounds present in drugs and environmental pollutants such as chlorobenzene, chlorocyclohexane, chloroalkane, and chloroalkene [62]. This enzyme presents applications in chemical industries, bioremediation, and sustainable chemistry [60, 61]. Results obtained for the pan-genome are shown in the additional file Table S4.

Genome analysis indicates that an efficient system for arsenic detoxification is restricted to *L. plantarum* WCFS1. This mechanism is regulated by the *arsR* gene and depends on ArsD, ArsA, and ArsB proteins [12, 65]. The other strains

contain only *arsC* and *arsR* genes, and therefore, they have the arsenic partial detoxification [64].

The nisin (*nsr*) gene is found in all strains analyzed in this study but encodes a protein truncated in ten strains (16, 5-2, JDM1, MF1298, p-8, ST-III, TMW 1.25, TMW 1.277, WCFS1, and ZJ316). In a similar way, Sun et al. [66] described in *L. lactis* a truncated nisin protein, with the activity reduced. Hence, it is expected that these strains also show a reduced nisin activity.

The production of vitamins in the food industry is greatly exploited by food biotechnology. However, *L. plantarum* is deficient in the production of vitamin B complex biotin (B7), niacin (B3), pantothenate (B5), and pyridoxine (B6) [67, 68]. But it produces large amounts of folate (B9) [69, 70]. Folate gene clusters are described by Kleerebezem et al. [12] in the *L. plantarum* WCFS1 genome. This cluster presents nine genes (*folA*, *folB*, *folC1*, *folC2*, *folD*, *folE*, *folK*, *folP*, and *folQ*) identified in the core genome of all strains analyzed here. In contrast, the presence of the ribo genes (*ribA*, *ribB*, *ribD*, and *ribH*) required for riboflavin synthesis

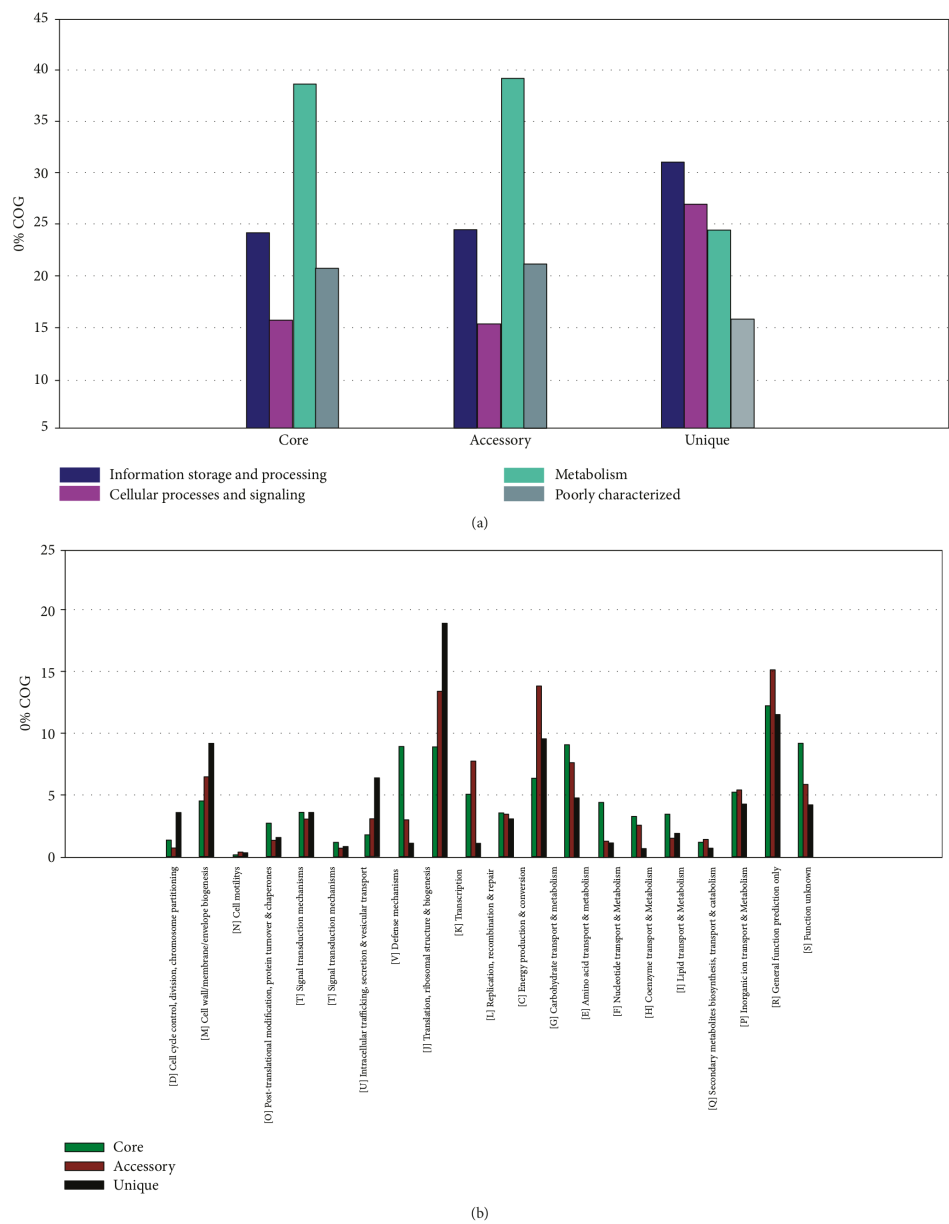


FIGURE 2: Pan-genome prediction. (a) core, accessory and unique genes into the functional standard of the COG; (b) the specific distribution of the genes into 20 COG categories.

is restricted to ten strains (the ATCC 8014, JDM1, KC28, LPL-1, TMW 1.25, TMW 1.277, TMW 1.708, TMW 1.1623, X7021, and ST-III strains) originally isolated from food. Other strains present a rib operon incomplete and probably do not produce riboflavin [71].

Besides the production of nisin and vitamins, the monitoring of the production of biogenic amines (BA) by LAB is also of paramount importance to the food industry. BA, such as putrescine and spermidine, are nitrogen compounds formed during the decarboxylation of amino acids by bacteria [72]. They are toxic when accumulated in food processing and storage, causing human health problems [73, 74]. In this way, the ability to produce large amounts of BA may be an obstacle to the use of some LAB. Alan et al. [75] monitored the ability of *L. plantarum* JDM1 to produce metabolites based on the decarboxylase test and its genic content. The authors concluded that the presence of the glutamate decarboxylase (*gadB*) gene is not enough to produce BA.

In this work, we observed that the gene encoding in the enzyme glutamate decarboxylase is common to all the strains analyzed. Another decarboxylase gene, *panD* (encoded 1-decarboxylase aspartate), was also found in the *L. plantarum* B21, *L. plantarum* TMW 1.708, and *L. plantarum* WCFS1 genomes. Based on genomic analysis, it is not possible to assess whether the amounts of BA produced are deleterious, but analyses in culture medium show that they are not [75, 76].

Only one plantaricin (*pln*) gene was identified in the core genome. This gene encodes a bacteriocin immunity protein with 88 amino acids (lp_2952 in reference genome WCFS1). Other genes are restricted to the accessory genome and unique gene families. The accessory genome has an average of 1320 OGs, mostly related to the phosphotransferase system (PTS) and biosynthesis of amino acids. PTS proteins transport substance into the cell, including carbohydrates. The sugar-specific transport of these proteins explains their greater genetic representation within the accessory genome of the *L. plantarum* strains.

A protein-coding gene involved in the export of bacteriocins, the bacteriocin ABC-transporter gene was found in most samples (except in the Asian strains GB-LP1, JBE490, LPL-1, LZ206, LZ227, Zhang-LL, and ZJ316). The mannose-specific adhesin (*msa*) gene also belongs to the accessory genome, being present in 22 strains extracted from different sources (fermented foods, flies, saliva, cell culture, faeces, environment, and probiotics). In addition, the collagen binding protein (*cnaB*) gene encodes an adhesion, likely related to colonisation and competition against pathogenic bacteria, an important feature of probiotic strains [77–79].

Phylogenetic trees obtained from the pan-genome and core genome are shown in Figures 3(a) and 3(b), respectively. The core genome tree recovered better the phylogenetic relationships between the strains (reference NCBI-Genome Tree Report: <https://www.ncbi.nlm.nih.gov/genome/tree/1108?>). In contrast, the pan-genome tree may show genomic novelties, such as the gaining of new genes by HGT [80, 81]. The monophyly of the strains extracted from flies and from potential probiotics was recovered in the phylogenetic trees. Two branches highlighted in Figure 3(a) indicate strains

grouped according to the geographical location where they were isolated. *L. plantarum* LZ206, ZJ316, and LZ227 strains are from Hangzhou, China, and have the same GC content (45.2%). The milk strains LZ206 and LZ227 share a CRISPR spacer (AAACGTTCTATGCTTCGTTTCCTCAGCATC) and are also the final part of a 74.2 kb foreign fragment. This may suggest a shift of genetic material between them. The origin of the cluster formed by *L. plantarum* TMW 1.25 + TMW 1.277 (monophyletic group), TMW 1.1623, and TMW 1.708 strains from Germany appears to be more complex than those of the other group. It has a similar CG content (between 45.2 and 45.4%) but does not have CRISPR sequences that could indicate recent invasions. Based on bacteriophage analysis, we were able to identify that the *L. plantarum* TMW 1.1623 strain partially shares with *L. plantarum* TMW 1.708 a 44.2 kb fragment (on positions 1129520 to 1173731), mainly containing the bacteriophage Lactob_Sha1 (35) and Lactob_JCL1032 (8), while with the *L. plantarum* TMW 1.25 and TMW 1.277 strains, it shares in a region of about 47.3 kb (positions 2074572–2121906), composed mainly of bacteriophages of the types Oenoco_phiS13 (16), Oenoco_phi9805 (15), and Lactob_Lj965 (14).

3.3. Bacteriocin Genes. Twenty-one strains present EF and JK plantaricin genes, which make up the *pln* cluster (formed by 25 genes). However, *L. plantarum* 16, *L. plantarum* C410L1, and *L. plantarum* subsp. *plantarum* p-8 strains have a frameshift in the *plnE* gene, and thus, the synthesis of EF plantaricin by them is questionable. These strains contain IS3 and IS256 insertion elements (IS) in the middle of the *pln* cluster, suggesting that IS may be related to the loss of function in the *plnE* gene. The bacteriocin genes present in the strains are shown in Figure 3(b).

Capy et al. [82], Schneider and Lenski [83], and Eraclio et al. [84] proposed that IS could have an adaptive function and play a significant role in the chromosomal rearrangement. This assertion is likely persuasive since IS and transposable elements can inactivate, insert, delete, or displace operons and gene cassettes, shifting the adaptive value of the microorganism within its habitat.

Protein-coding genes for class IIb bacteriocin, a lactobin A/cerein 7B protein, are restricted to CLP0611, JBE245, LPL-1, LY-78, Z227, and CGMCC 1.557 strains. PLNC8 $\alpha\beta$ genes expressed in *L. plantarum* NC8 [85] were also found in the genomes of MF1298 (fermented sausage, Norway) and LZ206 (cow milk, China) strains. PLNC8 $\alpha\beta$ proved effective in controlling *Porphyromonas gingivalis*, a bacterium that causes periodontitis [85, 86]. Thus, products based on the LZ206 and MF1298 strains may be potentially useful for the treatment of periodontitis. LPL-1 presents pediocin PA-1 bacteriocins (class IIa), similar to that found in "*Pediococcus acidilactici* H".

The origin of some bacteriocins is attributed to defective bacteriophage proteins, such as the R-type pyocin related to the P2 bacteriophage, carotovoricin to tail-like bacteriocin, and monocins to TP901-1-like bacteriophage tails [87–90]. We found no evidence that bacteriocins are bacteriophage-derived proteins; however, these proteins may be important in rearrangement and environmental adaptation [72–74].

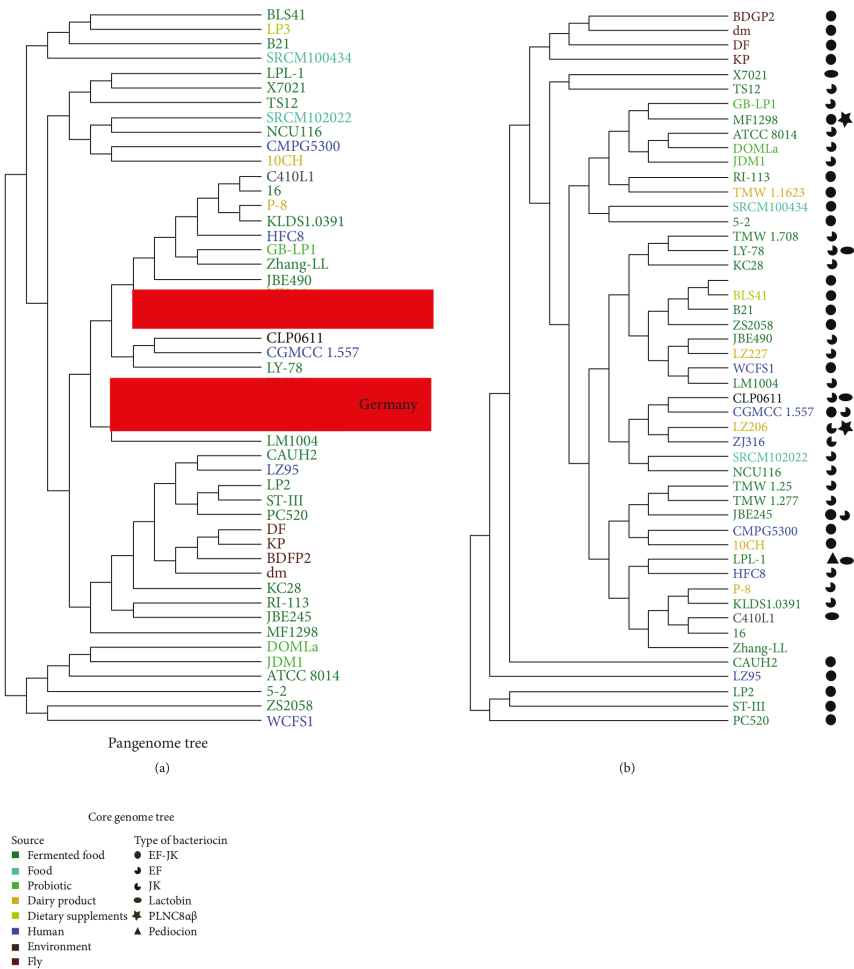


FIGURE 3: Phylogenies based on pangenome analyses of *L. plantarum*. (a) A phylogenetic tree constructed by pangenome data. The branches highlighted indicate the geographical locations where the strains were isolated. (b) A phylogenetic tree based on core genome data. The first box shows which colors correspond to the source of the strains, and the box below described the symbols used here to represent the bacteriocins. The scales below the trees correspond to the time in millions of years (Mya). Nisin bacteriocin is not shown since it is present in most strains.

3.4. *Multiple Genome Alignments.* The genomic arrangement possesses small variations among strains, mainly in the life-style island (Figure 4). We also conducted a comparative analysis using the lifestyle island (lp_3131 to lp_3661 position genes, using *L. plantarum* WCFS1 as reference) [3, 4, 23, 91]. The analysis of this genomic region did not show a pattern associated with the habitat of the strains (Figure 5). Some arrangements were consistent with the phylogenetic

relationship shown in Figure 3(b), for instance, *L. plantarum* ATCC 8014, DOMLa, and JDM1 strains, while other similar arrangements arose via HGT.

4. Conclusions

L. plantarum strains are potentially useful in biotechnology, bioremediation, and pharmaceutical products and in the

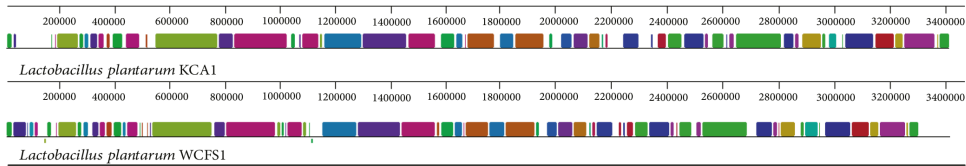


FIGURE 4: Genome comparison between *L. plantarum* WCFS1 and *L. pentosus* KCA1. The colored blocks correspond to LCBs.

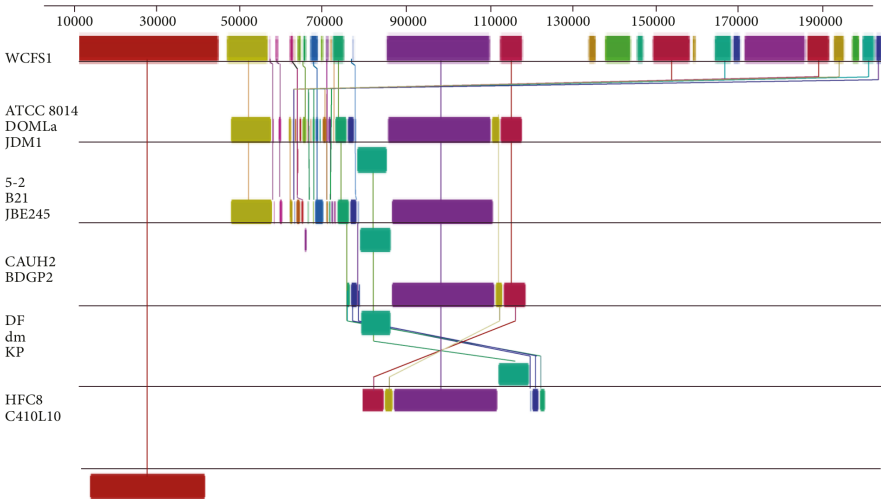


FIGURE 5: Comparative patterns of the lifestyle island in *L. plantarum* strains. The boxes below or above on the LCBs mean that the gene clusters are associated with obtaining energy. LCB weight value was 1000 bp. The region was analyzed corresponding to lp_3131 to lp_3661 of the *L. plantarum* WSFS1 genome.

manufacture of bacteriocins. None of the strains has antibiotic resistance genes or virulence factors. But the genomic screening of new strains is essential because the bacterial genomes are dynamic entities. HGT seems to play a large role on genomic innovations, and it may be related to the great adaptability of the *L. plantarum* to different ecological niches. In contrast, we found no evidence on the adaptive role of the lifestyle island.

Abbreviations

- arsC gene: Arsenate reductase gene
- arsR gene: Arsenical resistance operon transcriptional repressor gene
- BA: Biogenic amines
- cnaB gene: Collagen binding protein gene
- deh gene: L-2-haloacid dehalogenase gene
- HGT: Horizontal gene transfer
- HPK: Histidine protein kinase
- LCBs: Locally collinear blocks

- msa gene: Mannose-specific adhesin gene
- nsr gene: Nisin gene
- pln gene: Plantaricin gene
- srtA gene: Sortase A gene.

Data Availability

The supporting data are enclosed as additional files.

Additional Points

Authors' Information. This manuscript is part of the Ph.D. thesis of Eliane Evanovich from Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Federal do Pará.

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' Contributions

The authors read and approved the final manuscript.

Supplementary Materials

Supplementary 1. Table S1: the table shows the description of the samples used in the manuscript.

Supplementary 2. Table S2: the characteristics of the phages found in the genomes.

Supplementary 3. CRISPR sequences present in the analyzed genomes.

Supplementary 4. Detailed information on the pan-genome analysis.

References

- [1] S. Torriani, F. Clementi, M. Vancanneyt, B. Hoste, F. Dellaglio, and K. Kersters, "Differentiation of *Lactobacillus plantarum*, *L. pentosus* and *L. paraplantarum* species by RAPD-PCR and AFLP," *Systematic and Applied Microbiology*, vol. 24, no. 4, pp. 554–560, 2001.
- [2] E. A. Pfeiler and T. R. Klaenhammer, "The genomics of lactic acid bacteria," *Trends in Microbiology*, vol. 15, no. 12, pp. 546–553, 2007.
- [3] R. J. Siezen, V. A. Tzeneva, A. Castioni et al., "Phenotypic and genomic diversity of *Lactobacillus plantarum* strains isolated from various environmental niches," *Environmental Microbiology*, vol. 12, no. 3, pp. 758–773, 2010.
- [4] R. J. Siezen and J. E. T. van Hylckama Vlieg, "Genomic diversity and versatility of *Lactobacillus plantarum*, a natural metabolic engineer," *Microbial Cell Factories*, vol. 10, article S3, Supplement 1, 2011.
- [5] J. Wang, H. Ji, D. Zhang et al., "Assessment of probiotic properties of *Lactobacillus plantarum* ZLP001 isolated from gastrointestinal tract of weaning pigs," *African Journal of Biotechnology*, vol. 10, no. 54, pp. 11303–11308, 2011.
- [6] Y. Nami, N. Abdullah, B. Haghsheenas, D. Radiah, R. Rosli, and A. Y. Khosroushahi, "Assessment of probiotic potential and anticancer activity of newly isolated vaginal bacterium *Lactobacillus plantarum* 5BL," *Microbiology and Immunology*, vol. 58, no. 9, pp. 492–502, 2014.
- [7] R. Tabasco, F. Sánchez-Patán, M. Monagas et al., "Effect of grape polyphenols on lactic acid bacteria and Bifidobacteria growth: resistance and metabolism," *Food Microbiology*, vol. 28, no. 7, pp. 1345–1352, 2011.
- [8] P. Fras, F. M. Campos, T. Hogg, and J. A. Couto, "Production of volatile phenols by *Lactobacillus plantarum* in wine conditions," *Biotechnology Letters*, vol. 36, no. 2, pp. 281–285, 2014.
- [9] E. J. Yang and H. C. Chang, "Purification of a new antifungal compound produced by *Lactobacillus plantarum* AF1 isolated from kimchi," *International Journal of Food Microbiology*, vol. 139, no. 1–2, pp. 56–63, 2010.
- [10] P. Li, Q. Zhou, and Q. Gu, "Complete genome sequence of *Lactobacillus plantarum* LZ227, a potential probiotic strain producing B-group vitamins," *Journal of Biotechnology*, vol. 234, pp. 66–70, 2016.
- [11] K. B. Ahn, J. B. Baik, O. Park, C. Yun, and S. H. Han, "*Lactobacillus plantarum* lipoteichoic acid inhibits biofilm formation of *Streptococcus mutans*," *PLoS One*, vol. 13, no. 2, article e0192694, 2018.
- [12] M. Kleerebezem, J. Boekhorst, R. van Kranenburg et al., "Complete genome sequence of *Lactobacillus plantarum* WCFS1," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 4, pp. 1990–1995, 2003.
- [13] J. C. Hubert and B. Kammerer, "Determination of chromosome size and number of *rrn* loci in *Lactobacillus plantarum* by pulsed-field gel electrophoresis," *FEMS Microbiology Letters*, vol. 120, no. 1–2, pp. 51–56, 1994.
- [14] J. M. Neal-McKinney, X. Lu, T. Duong et al., "Production of organic acids by probiotic lactobacilli can be used to reduce pathogen load in poultry," *PLoS One*, vol. 7, no. 9, article e43928, 2012.
- [15] O. Cortés-Zavaleta, A. López-Malo, A. Hernández-Mendoza, and H. S. García, "Antifungal activity of lactobacilli and its relationship with 3-phenyllactic acid production," *International Journal of Food Microbiology*, vol. 173, pp. 30–35, 2014.
- [16] A. Guimarães, A. Santiago, J. A. Teixeira, A. Venâncio, and L. Abrunhosa, "Anti-aflatoxigenic effect of organic acids produced by *Lactobacillus plantarum*," *International Journal of Food Microbiology*, vol. 264, pp. 31–38, 2018.
- [17] P. D. Cotter, C. Hill, and R. P. Ross, "Bacteriocins: developing innate immunity for food," *Nature Reviews. Microbiology*, vol. 3, no. 10, pp. 777–788, 2005.
- [18] F. Atassi and A. L. Servin, "Individual and co-operative roles of lactic acid and hydrogen peroxide in the killing activity of enteric strain *Lactobacillus johnsonii* NCC933 and vaginal strain *Lactobacillus gasseri* KS120.1 against enteric, uropathogenic and vaginosis-associated pathogens," *FEMS Microbiology Letters*, vol. 304, no. 1, pp. 29–38, 2010.
- [19] D. L. Turner, L. Brennan, H. E. Meyer et al., "Solution structure of plantaricin C, a novel lantibiotic," *European journal of biochemistry*, vol. 264, no. 3, pp. 833–839, 1999.
- [20] H. Zhang, L. Liu, Y. Hao et al., "Isolation and partial characterization of a bacteriocin produced by *Lactobacillus plantarum* BM-1 isolated from a traditionally fermented Chinese meat product," *Microbiology and Immunology*, vol. 57, no. 11, pp. 746–755, 2013.
- [21] H. A. Seddik, F. Bendali, F. Gancel, I. Fliss, G. Spano, and D. Drider, "*Lactobacillus plantarum* and Its probiotic and food potentialities," *Probiotics Antimicro Prot*, vol. 9, no. 2, pp. 111–122, 2017.
- [22] D. Molenaar, F. Bringel, F. H. Schuren, W. M. de Vos, R. J. Siezen, and M. Kleerebezem, "Exploring *Lactobacillus plantarum* genome diversity by using microarrays," *Journal of Bacteriology*, vol. 187, no. 17, pp. 6119–6127, 2005.
- [23] M. N. Alekshun and S. B. Levy, "Molecular mechanisms of antibacterial multidrug resistance," *Cell*, vol. 128, no. 6, pp. 1037–1050, 2007.
- [24] M. T. Holden, E. J. Feil, J. A. Lindsay et al., "Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, pp. 9786–9791, 2004.
- [25] S. R. Partridge, G. Tsafnat, E. Coiera, and J. R. Iredell, "Gene cassettes and cassette arrays in mobile resistance integrons," *FEMS Microbiology Reviews*, vol. 33, no. 4, pp. 757–784, 2009.

- [26] M. Munoz-Lopez and J. L. Garcia-Perez, "DNA transposons: nature and applications in genomics," *Current Genomics*, vol. 11, no. 2, pp. 115–128, 2010.
- [27] T. Zhang, X. Zhang, and L. Ye, "Plasmid metagenome reveals high levels of antibiotic resistance genes and mobile genetic elements in activated sludge," *PLoS One*, vol. 6, no. 10, article e26041, 2011.
- [28] O. Cataloluk and B. Gogebakan, "Presence of drug resistance in intestinal lactobacilli of dairy and human origin in Turkey," *FEMS Microbiology Letters*, vol. 236, no. 1, pp. 7–12, 2004.
- [29] S. Kastner, V. Perreten, H. Bleuler, G. Hugenschmidt, C. Lacroix, and L. Meile, "Antibiotic susceptibility patterns and resistance genes of starter cultures and probiotic bacteria used in food," *Systematic and Applied Microbiology*, vol. 29, no. 2, pp. 145–155, 2006.
- [30] L. Drago, V. Rodighiero, R. Mattina, M. Toscano, and E. DE Vecchi, "In vitro selection and transferability of antibiotic resistance in the probiotic strain *Lactobacillus reuteri* DSM 17938," *Journal of Chemotherapy*, vol. 23, no. 6, pp. 371–373, 2011.
- [31] A. B. Flórez, M. Egerbærn, M. Danielsen et al., "Susceptibility of *Lactobacillus plantarum* strains to six antibiotics and definition of new susceptibility resistance cutoff values," *Microbial Drug Resistance*, vol. 12, no. 4, pp. 252–256, 2006.
- [32] M. Danielsen, "Characterization of the tetracycline resistance plasmid pMD5057 from *Lactobacillus plantarum* 5057 reveals a composite structure," *Plasmid*, vol. 48, no. 2, pp. 98–103, 2002.
- [33] G. Huys, K. D'Haene, and J. Swings, "Genetic basis of tetracycline and minocycline resistance in potentially probiotic *Lactobacillus plantarum* strain CCUG 43738," *Antimicrobial Agents and Chemotherapy*, vol. 50, no. 4, pp. 1550–1551, 2006.
- [34] L. Jacobsen, A. Wilcks, K. Hammer, G. Huys, D. Gevers, and S. R. Andersen, "Horizontal transfer of tet(M) and erm(B) resistance plasmids from food strains of *Lactobacillus plantarum* to *Enterococcus faecalis* JH2-2 in the gastrointestinal tract of gnotobiotic rats," *FEMS Microbiology Ecology*, vol. 59, no. 1, pp. 158–166, 2007.
- [35] E. Branscomb and P. Predki, "On the high value of low standards," *Journal of Bacteriology*, vol. 184, no. 23, pp. 6406–6409, 2002.
- [36] S. N. Gardner, M. W. Lam, J. R. Smith, C. T. Torres, and T. R. Slezak, "Draft versus finished sequence data for DNA and protein diagnostic signature development," *Nucleic Acids Research*, vol. 33, no. 18, pp. 5838–5850, 2005.
- [37] B. H. Yoon, S. H. Jang, and H. Chang, "Sequence analysis of the *Lactobacillus temperate* phage Sha1," *Archives of Virology*, vol. 156, no. 9, pp. 1681–1684, 2011.
- [38] K. I. Kodaira, M. Oki, M. Kakikawa et al., "Genome structure of the *Lactobacillus temperate* phage ϕ g1e: the whole genome sequence and the putative promoter/repressor system," *Gene*, vol. 187, no. 1, pp. 45–53, 1997.
- [39] M. Kakikawa, M. Oki, H. Tadokoro, S. Nakamura, A. Taketo, and K. Kodaira, "Cloning and nucleotide sequence of the major capsid proteins of *Lactobacillus* bacteriophage Φ gle," *Gene*, vol. 175, no. 1–2, pp. 157–165, 1996.
- [40] T. Seemann, "Prokka: rapid prokaryotic genome annotation," *Bioinformatics*, vol. 30, no. 14, pp. 2068–2069, 2014.
- [41] I. Grissa, G. Vergnaud, and C. Pourcel, "CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats," *Nucleic Acids Research*, vol. 35, 2, pp. W52–W57, 2007.
- [42] D. Arndt, A. Marcu, Y. Liang, and D. S. Wishart, "PHAST, PHASTER and PHASTEST: tools for finding prophage in bacterial genomes," *Briefings in Bioinformatics*, no. article bbx121, 2017.
- [43] B. Jia, A. R. Raphenya, B. Alcock et al., "CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database," *Nucleic Acids Research*, vol. 45, no. D1, pp. D566–D573, 2017.
- [44] N. M. Chaudhari, V. K. Gupta, and C. Dutta, "BPGA- an ultra-fast pan-genome analysis pipeline," *Scientific Reports*, vol. 6, no. 1, p. 24373, 2016.
- [45] R. C. Edgar, "Search and clustering orders of magnitude faster than BLAST," *Bioinformatics*, vol. 26, no. 19, pp. 2460–2461, 2010.
- [46] R. C. Edgar, "MUSCLE: multiple sequence alignment with high accuracy and high throughput," *Nucleic Acids Research*, vol. 32, no. 5, pp. 1792–1797, 2004.
- [47] T. Williams and C. Kelley, "Gnuplot 4.6.6. An interactive plotting program," 2014, March 2018, http://www.gnuplot.info/docs_4.6/gnuplot.pdf.
- [48] A. E. Darling, B. Mau, F. R. Blattner, and N. T. Perna, "Mauve: multiple alignment of conserved genomic sequence with rearrangements," *Genome Research*, vol. 14, no. 7, pp. 1394–1403, 2004.
- [49] A. E. Darling, B. Mau, and N. T. Perna, "ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement," *PLoS One*, vol. 5, no. 6, article e11147, 2010.
- [50] G. Gasiunas, R. Barrangou, P. Horvath, and V. Siksnys, "Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 39, pp. E2579–E2586, 2012.
- [51] H. Abriouel, B. Pérez Montoro, M. Casado Muñoz, C. W. Knapp, A. Gálvez, and N. Benomar, "In silico genomic insights into aspects of food safety and defense mechanisms of a potentially probiotic *Lactobacillus pentosus* MP-10 isolated from brines of naturally fermented Alor-eña green table olives," *PLoS One*, vol. 12, no. 6, article e0176801, 2017.
- [52] K. C. Anukam, J. M. Macklaim, G. B. Gloor, G. Reid, and J. Boekhorst, "Genome sequence of *Lactobacillus pentosus* KCA1: vaginal isolate from a healthy premenopausal woman," *PLoS One*, vol. 8, no. 3, article e59239, 2013.
- [53] L. M. Bobay, M. Touchon, and E. P. Rocha, "Pervasive domestication of defective prophages by bacteria," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 33, pp. 12127–12132, 2014.
- [54] G. Edlin, L. Lin, and R. Bitner, "Reproductive fitness of P1, P2, and Mu lysogens of *Escherichia coli*," *Journal of Virology*, vol. 21, no. 2, pp. 560–564, 1977.
- [55] J. Gödeke, K. Paul, J. Lassak, and K. M. Thormann, "Phage-induced lysis enhances biofilm formation in *Shewanella oneidensis* MR-1," *The ISME Journal*, vol. 5, no. 4, pp. 613–626, 2011.
- [56] X. Wang, Y. Kim, Q. Ma et al., "Cryptic prophages help bacteria cope with adverse environments," *Nature Communications*, vol. 1, no. 9, p. 147, 2010.

- [57] M. E. Martino, J. R. Bayjanov, B. E. Caffrey et al., "Nomadic lifestyle of *Lactobacillus plantarum* revealed by comparative genomics of 54 strains isolated from different habitats," *Environmental Microbiology*, vol. 18, no. 12, pp. 4974–4989, 2016.
- [58] G. S. Cho, M. Huch, A. Hanak, W. H. Holzapfel, and C. M. Franz, "Genetic analysis of the plantaricin EFI locus of *Lactobacillus plantarum* PCS20 reveals an unusual plantaricin E gene sequence as a result of mutation," *International Journal of Food Microbiology*, vol. 141, pp. S117–S124, 2010.
- [59] G. Dirix, P. Monsieurs, B. Dombrecht et al., "Peptide signal molecules and bacteriocins in gram-negative bacteria: a genome-wide in silico screening for peptides containing a double-glycine leader sequence and their cognate transporters," *Peptides*, vol. 25, no. 9, pp. 1425–1440, 2004.
- [60] T. Kurihara and N. Esaki, "Bacterial hydrolytic dehalogenases and related enzymes: occurrences, reaction mechanisms, and applications," *Chemical Record*, vol. 8, no. 2, pp. 67–74, 2008.
- [61] K. Furukawa, "Oxygenases and dehalogenases: molecular approaches to efficient degradation of chlorinated environmental pollutants," *Bioscience, Biotechnology, and Biochemistry*, vol. 70, no. 10, pp. 2335–2348, 2006.
- [62] Y. Wang, X. Cao, Y. Feng, and S. Xue, "Environment-induced conformational and functional changes of 1-2-haloacid dehalogenase," *Journal of Bioscience and Bioengineering*, vol. 121, no. 5, pp. 491–496, 2016.
- [63] Y. Wang, Y. Feng, X. Cao, Y. Liu, and S. Xue, "Insights into the molecular mechanism of dehalogenation catalyzed by D-2-haloacid dehalogenase from crystal structures," *Scientific Reports*, vol. 8, no. 1, p. 1454, 2018.
- [64] B. P. Rosen, "Biochemistry of arsenic detoxification," *FEBS Letters*, vol. 529, no. 1, pp. 86–92, 2002.
- [65] R. Van Kranenburg, N. Golic, R. Bongers et al., "Functional analysis of three plasmids from *Lactobacillus plantarum*," *Applied and Environmental Microbiology*, vol. 71, no. 3, pp. 1223–1230, 2005.
- [66] Z. Sun, J. Zhong, X. Liang, J. Liu, X. Chen, and L. Huan, "Novel mechanism for nisin resistance via proteolytic degradation of nisin by the nisin resistance protein NSR," *Antimicrobial Agents and Chemotherapy*, vol. 53, no. 5, pp. 1964–1973, 2009.
- [67] O. Kandler and N. Weiss, "Genus *Lactobacillus* Beijerinck 1901, 212AL," in *Bergey's Manual of Systematic Bacteriology*, P. H. A. Sneath, N. S. Mair, M. E. Sharpe, and J. G. Holt, Eds., vol. 2, pp. 1209–1234, Williams & Wilkins, Baltimore, 1986.
- [68] O. V. Ledesma, A. P. De Ruiz Holgado, G. Oliver, G. S. De Giori, P. Raibaud, and J. V. Galpin, "A synthetic medium for comparative nutritional studies of *Lactobacilli*," *Journal of Applied Bacteriology*, vol. 42, no. 1, pp. 123–133, 1977.
- [69] A. Wegkamp, A. E. Mars, M. Fajies et al., "Physiological responses to folate overproduction in *Lactobacillus plantarum* WCFS1," *Microbial Cell Factories*, vol. 9, no. 1, p. 100, 2010.
- [70] M. Masuda, M. Ide, H. Utsumi, T. Niino, Y. Shimamura, and M. Murata, "Production potency of folate, vitamin B₁₂, and thiamine by lactic acid bacteria isolated from Japanese pickles," *Bioscience, Biotechnology, and Biochemistry*, vol. 76, no. 11, pp. 2061–2067, 2012.
- [71] C. Burgess, M. O'Connell-Motherway, W. Sybesma, J. Hugenholtz, and D. van Sinderen, "Riboflavin production in *Lactococcus lactis*: potential for in situ production of vitamin-enriched foods," *Applied and Environmental Microbiology*, vol. 70, no. 10, pp. 5769–5777, 2004.
- [72] M. V. Moreno-Arribas, M. C. Polo, F. Jorganes, and R. Munoz, "Screening of biogenic amine production by lactic acid bacteria isolated from grape must and wine," *International Journal of Food Microbiology*, vol. 84, no. 1, pp. 117–123, 2003.
- [73] V. Capozzi, P. Russo, V. Ladero et al., "Biogenic amines degradation by *Lactobacillus plantarum*: toward a potential application in wine," *Frontiers in Microbiology*, vol. 3, p. 122, 2012.
- [74] Z. Y. Zhang, C. Liu, Y. Z. Zhu et al., "Safety assessment of *Lactobacillus plantarum* JDM1 based on the complete genome," *International Journal of Food Microbiology*, vol. 153, no. 1–2, pp. 166–170, 2012.
- [75] Y. Alan, Z. Topalcengiz, and M. Dıgırak, "Biogenic amine and fermentation metabolite production assessments of *Lactobacillus plantarum* isolates for naturally fermented pickles," *LWT - Food Science and Technology*, vol. 98, pp. 322–328, 2018.
- [76] Y. C. Lee, H. F. Kung, Y. L. Huang, C. H. Wu, Y. R. Huang, and Y. H. Tsai, "Reduction of biogenic amines during miso fermentation by *Lactobacillus plantarum* as a starter culture," *Journal of Food Protection*, vol. 79, no. 9, pp. 1556–1561, 2016.
- [77] A. K. Yadav, A. Tyagi, J. K. Kaushik, A. C. Saklani, S. Grover, and V. K. Batish, "Role of surface layer collagen binding protein from indigenous *Lactobacillus plantarum* 91 in adhesion and its anti-adhesion potential against gut pathogen," *Microbiological Research*, vol. 168, no. 10, pp. 639–645, 2013.
- [78] D. Muñoz-Provencio, G. Pérez-Martínez, and V. Monedero, "Characterization of a fibronectin-binding protein from *Lactobacillus casei* BL23," *Journal of Applied Microbiology*, vol. 108, no. 3, pp. 1050–1059, 2010.
- [79] D. Muñoz-Provencio and V. Monedero, "Shotgun phage display of *Lactobacillus casei* BL23 against collagen and fibronectin," *Journal of Microbiology and Biotechnology*, vol. 21, no. 2, pp. 197–203, 2011.
- [80] E. Mira, A. B. Martín-Cuadeado, G. D'Auria, and F. Rodríguez-Valera, "The bacterial pan-genome: a new paradigm in microbiology," *International Microbiology*, vol. 13, pp. 45–57, 2010.
- [81] L. D. Alcaraz, "Pan-genomics: unmasking the gene diversity hidden in the bacteria species," *PeerJ PrePrints*, vol. 2, 2014.
- [82] P. Capy, G. Gasperi, C. Biémont, and C. Bazin, "Stress and transposable elements: co-evolution or useful parasites?," *Heredity*, vol. 85, no. 2, pp. 101–106, 2000.
- [83] D. Schneider and R. E. Lenski, "Dynamics of insertion sequence elements during experimental evolution of bacteria," *Research in Microbiology*, vol. 155, no. 5, pp. 319–327, 2004.
- [84] G. Eracilo, G. Ricci, and M. G. Fortina, "Insertion sequence elements in *Lactococcus garvieae*," *Gene*, vol. 555, no. 2, pp. 291–296, 2015.
- [85] H. Khalaf, S. S. Nakka, C. Sandén et al., "Antibacterial effects of *Lactobacillus* and bacteriocin PLNC8 $\alpha\beta$ on the periodontal pathogen *Porphyromonas gingivalis*," *BMC Microbiology*, vol. 16, no. 1, p. 188, 2016.
- [86] A. Maldonado, J. L. Ruiz-Barba, and R. Jimenez-Diaz, "Purification and genetic characterization of plantaricin NC8, a novel coculture-inducible two-peptide bacteriocin from *Lactobacillus plantarum* NC8," *Applied and Environmental Microbiology*, vol. 69, no. 1, pp. 383–389, 2003.

- [87] S. J. Labrie, J. Samson, and S. Moineau, "Bacteriophage resistance mechanisms," *Nature Reviews Microbiology*, vol. 8, no. 5, pp. 317–327, 2010.
- [88] K. Nakayama, K. Takashima, H. Ishihara et al., "The R-type pyocin of *Pseudomonas aeruginosa* is related to P2 phage, and the F-type is related to lambda phage," *Molecular Microbiology*, vol. 38, no. 2, pp. 213–231, 2000.
- [89] K. Yamada, M. Hirota, Y. Niimi et al., "Nucleotide sequences and organization of the genes for carotovoricin (Ctv) from *Erwinia carotovora* indicate that Ctv evolved from the same ancestor as *Salmonella typhi* prophage," *Bioscience, Biotechnology, and Biochemistry*, vol. 70, no. 9, pp. 2236–2247, 2006.
- [90] G. Lee, U. Chakraborty, D. Gebhart, G. R. Govoni, Z. H. Zhou, and D. Scholl, "F-type bacteriocins of *Listeria monocytogenes*: a new class of phage tail-like structures reveals broad parallel coevolution between tailed bacteriophages and high-molecular-weight bacteriocins," *Journal of Bacteriology*, vol. 198, no. 20, pp. 2784–2793, 2016.
- [91] M. Van den Nieuwboer, S. van Hemert, E. Claassen, and W. M. de Vos, "Lactobacillus plantarum WCFS1 and its host interaction: a dozen years after the genome," *Microbial Biotechnology*, vol. 9, no. 4, pp. 452–465, 2016.

4. DISCUSSÃO

Bacteroides fragilis

Bacteroides fragilis é um bacilo que pode causar infecção grave em hospedeiros suscetíveis e a presença de linhagens multirresistentes na infecção pode dificultar o tratamento. A obtenção de mais informações sobre os genomas dessas bactérias pode auxiliar no desenvolvimento de novas terapias para o seu controle. Desta forma, o estudo desta bactéria é essencial até mesmo para maior entendimento do desenvolvimento de câncer colorretal, visto que sua toxina, a fragilisina é associada a esta doença.

Nossos resultados indicaram que o estudo de *contigs* e *scaffolds* também pode ser informativos, embora genomas completos gerem dados mais robustos. Sem dúvida, os genes de resistência a antibióticos e seus meios de disseminação apresentam mais destaque nos estudos genômicos destas bactérias. O principal veículo de disseminação destes genes parece ser o ICE com T4SS; no entanto, CTn341 e ISBf1 aparentemente tiveram grande relevância na propagação ancestral dos genes *tetQ* e *cepA*, mas muitos dos seus genes perderam suas funções. O transposon Ctn86 ajudou na disseminação do gene da fragilisina entre as linhagens e seus genes estão integros, indicando que esta função pode estar ainda ativa. Também encontramos evidências de HGT entre uma região da linhagem DCMOUH0085B, que abriga os genes de resistência *aadE* e *ermB* e uma região de *Butyricimonas faecalis* H184, ambos são ICE putativos com T4SS. Estas duas bactérias são Bacteroidales, sugerindo que HGT pode ser comum dentro da ordem, corroborando análises anteriores (como SHOEMAKER *et al.*, 1989; DELAVAT *et al.*, 2017; García-Bayona *et al.* (2021) que demonstram intensa troca de material genético entre essas bactérias.

O sistema CRISPR-Cas tipo IIIB é o mais frequente entre as linhagens enterotoxigênicas. Em contraste, a maioria das linhagens não enterotoxigênicas não possui um sistema CRISPR-Cas. Este mesmo padrão foi encontrado por Tajkarimi e Wexler (2017). Não conseguimos definir um motivo para este padrão, mas é possível que seja elucidado futuramente, através da análises usando mais genomas completos.

A relativa frequência de genes sob seleção positiva (396 genes) é coerente com a característica combativa destas bactérias, visto que elas apresentam sistemas e genes diferenciados que atuam contra bactérias antagônicas, tais como o sistema T6S e sistema bacteriano de dois componentes (*Bacterial two-component system* -TCS), descrito nestas bactérias por Wakimoto *et al.* (2013).

Lactiplantibacillus plantarum

L. plantarum é encontrada em diferentes habitats (plantas, alimentos fermentados, saliva, microbiota intestinal de vertebrados e invertebrados, etc) e é amplamente usada como probiótico para regularizar as defesas intestinais.

Conforme observado em nossas análises. *L. plantarum* apresenta características genômicas exclusivas, como a presença de bacteriocinas específicas e sintenia diferente de outros membros do gênero *Lactobacillus*. Desta forma, a mudança do nome do gênero *Lactobacillus* para *Lactiplantibacillus plantarum* proposta por Zheng *et al.*, (2020) é coerente.

Muitos genes destas bactérias estão relacionados à regulação, metabolismo e transporte de diferentes carboidratos. As análises das ilhas de estilo de vida não mostraram

um padrão de arranjo ou conteúdo genético de acordo com o habitat, como foi sugerido na literatura. Nossos resultados demonstraram que seus genomas contêm vários genes adquiridos por HGT, no entanto, não identificamos quaisquer fatores de virulência e genes de resistência a antibióticos. Desta forma, as linhagens podem ser úteis à biotecnologia, indústria de fármacos, biorremediação e produção de bacteriocinas. Mas, a triagem genômica de novas linhagens é essencial, porque os genomas bacterianos são entidades dinâmicas e HGT parece ter um grande papel nas inovações genômicas, podendo estar relacionado à grande adaptabilidade do *L. plantarum* a diferentes nichos ecológicos.

Contrastes Genômicos

Além de apresentarem comportamento diferentes na microbiota intestinal humana, *B. fragilis* e *L. plantarum* apresentam diferenças genômicas marcantes. Algumas delas são descritas abaixo.

B. fragilis apresenta genomas maiores e consequentemente com maior conteúdo gênico (em média 5,3 Mb e 4.332 genes ortólogos) que os descritos em *L. plantarum* (cerca de 3 Mb e 2919 genes ortólogos identificados). Mas a maioria dos genes, principalmente nos genomas centrais e acessórios, codificam proteínas envolvidas na síntese de carboidratos e aminoácidos, seguindo um padrão comum identificado em bactérias da microbiota intestinal humana (ZOU *et al.*, 2019). Outra semelhança é que os seus pan-genomas estão abertos, portanto, estão aptas a receber novos genes e sofrer adaptação. O sistema CRISPR-Cas é pouco diversificado em *L. plantarum* em comparação ao encontrado em *B. fragilis*. Além disso, o genoma de *B. fragilis* é um

reservatório de genes de resistência à antibióticos e genes de virulência que podem ser disseminados por diferentes vias.

Programas utilizados nas análises genômicas

Vários programas e pacotes de programas foram usados durante as análises genômicas. Como exposto, há várias opções para a realização destas análises, mas alguns critérios particulares foram essenciais para estipular e justificar o uso de cada programa. Esta exposição é feita abaixo:

- 1) Comparação genômica: O programa MAUVE consegue mostrar a comparação entre regiões genômicas de forma visualmente elegante e permite ainda a procura por genes específicos, portanto foi ideal para mostrar a sintenia entre as regiões genômicas nas duas espécies;
- 2) Análise dos Pan-genoma: O uso de pacote BPGA é justificado por vários motivos. O primeiro é que ele aceita sequências *drafts* como arquivo de entrada e gera arquivos de saída extremamente informativos que mostram filogenia de pan-genoma e do genoma central, além de apresentar as análises funcionais baseadas no COG e KEGG;
- 3) Predição das Sequências CRISPR-Cas: Até a publicação do artigo sobre *L. plantarum*, o programa CRISPRFinder era o mais indicado para a predição da sequências CRISPR-Cas, mas sua versão atualizada CRISPRCasFinder, apresenta mais funções que a versão anterior, portanto, foi usada nas análises de *B. fragilis*;
- 4) Predição das Sequências Profágicas: Durante essas análises somente genomas completos foram analisados. O Prophage Hunter foi utilizado na análise dos genomas de *L.*

plantarum, enquanto que o PhiSpy foi usado na análise de *B. fragilis*. A análise de Roach *et al.* (2021), demonstrou que o PhiSpy é o software que apresentou maior precisão na busca por sequências profágicas.

6) Predição de genes de resistência a antibiótico: Na análise de *L. planturum* usamos os programas CARD e ResFinder 4.0, mas o último foi rejeitado na análise de *B. fragilis*, pois este programa falhou em encontrar genes de resistência descritos anteriores em certas linhagens. O CARD fornece informações para uma análise bastante específica que integra vários outros bancos de dados, além de contar com as opções de análise pelo RGI e BLAST;

7) Identificação de Genes de Virulência: VFDB e Victor's são bancos de dados mais conservadores na busca por genes relacionados à virulência. Por sua vez, o banco PATRIC indicou muitos genes aparentemente falso positivos com baixos escores (inferiores a 30% de identidade) quando avaliados pelo BLAST-VFDB;

8) Análise de Evolução Molecular: O pacote Hyphy é atualmente a ferramenta mais completa para análise de evolução molecular, contando com a busca de rearranjos, a implementação de modelo evolutivo nas análises e diferentes métodos para a detecção de seleção positiva em ramos e sítios; além de detectar seleção episódica ou relaxada.

Impacto Acadêmico das Publicações

O artigo **Comparative genomics analysis of *Bacteroides fragilis*** ainda não foi publicado, portanto, não podemos avaliar seu impacto. No entanto, **Comparative Genomic Analysis of *Lactobacillus plantarum*: An Overview** foi publicada em 10 de

abril de 2019 e foi citada em dezenove publicações, treze destas no ano de 2021 (17 artigos e dois capítulos de livros). As publicações citadas são abaixo.

- 1) MISSAOUI, J.; SAIDANE, D.; MZOUGH, R.; MINERVINI, F. Fermented Seeds ("Zgougou") from Aleppo Pine as a Novel Source of Potentially Probiotic Lactic Acid Bacteria. **Microorganisms**, v. 7, n. 709, p. 1-23, dec, 2019, doi:10.3390/microorganisms7120709
- 2) WANG, Y.; JIANG, Y.; DEG, Y.; YI, C.; WANG, Y.; DING, M.; LIU, J.; JIN, X.; SHEN, L.; HE, Y.; WU, X.; CHEN, X.; SUN, C.; ZHENG, M.; ZHANG, R.; YE, H.; AN, H.; WONG, A. Probiotic Supplements: Hope or Hype? **Frontiers in Microbiology**, v. 11, n. 160, fev, 2020, doi:10.3389/fmicb.2020.00160
- 3) YAO, W.; YANG, L.; SHAO, Z.; XIE, L.; CHEN, L. Identification of salt tolerance-related genes of *Lactobacillus plantarum* D31 and T9 strains by genomic analysis. **Annals of Microbiology**, v. 70, n. 10, jan, 2020, doi: 10.1186/s13213-020-01551-2
- 4) KIM, E.; CHANG, H. C.; KIM, H. Y. Complete Genome Sequence of *Lactobacillus plantarum* EM, A Putative Probiotic Strain with the Cholesterol-Lowering Effect and Antimicrobial Activity. **Current Microbiology**, v. 77, n. 8, p. 1871-1882, ago, 2020, doi: 10.1007/s00284-020-02000-8
- 5) BACHTARZI, N.; SPECIALE, I.; KHARROUB, K.; DE CASTRO, C.; RUIZ, L.; RUAS-MADIEDO, P. Selection of Exopolysaccharide-Producing *Lactobacillus Plantarum* (*Lactiplantibacillus Plantarum*) Isolated from Algerian Fermented Foods for the Manufacture of Skim-Milk Fermented Products. **Microorganisms**, v. 8, n. 8, p. 1-21, jul, 2020, doi: 10.3390/microorganisms8081101
- 6) INDRAKUSUMA, A. A. B. P.; SADEVA, G. K. A.; WULANDARI, P. A.; WIGUNA, G. W. W.; REMITHA, N. P. S.; SADVIKA, G. A. S.; SUPADMANABA, G. P.; WIHANDANI, D. M. 2020. Potensi *Lactobacillus plantarum* sebagai intervensi penatalaksanaan kanker kolorektal berbasis modifikasi mikrobiota usus. **Jurnal Kesehatan**, v. 13, n. 2, p. 144-157, dez, 2020, doi: 10.24252/kesehatan.v13i2.17044
- 7) RODRÍGUEZ-LÓPEZ, C. M.; GUZMÁN-BELTRÁN, A. M.; LARA-MORALES, M. C.; CASTILLO, E.; BRANDÃO, P. F. B. Aislamiento e identificación de *Lactobacillus spp.* (LACTOBACILLACEAE) resistentes a Cd(II) y As(III) recuperados de fermento de cacao. **Acta Biológica Colombiana**, v. 26, n. 1, p. 19-29, mar, 2021, doi: 10.15446/abc.v26n1.83677
- 8) ALTERMANN, E.; BAILIE, M. A.; FRASER, K.; YOUNG, W. 2021. NexGen Sequencing Data: Bioinformatic Tools for Visualization and Analysis. Ed. CIFUENTES, A. In: **Comprehensive Foodomics**. Elsevier, Amsterdã, Países Baixos, 32 p, doi: 10.1016/b978-0-12-816395-5.00001-0

- 9) GARCIA-GONZALEZ, N.; BATTISTA, N.; PRETE, R., CORSETTI, A. 2021. Health-Promoting Role of *Lactiplantibacillus plantarum* Isolated from Fermented Foods. **Microorganisms** v. 9, n. 2, p. 2-30, fev, 2021, doi: 10.3390/microorganisms9020349
- 10) TENEA, G.N; ORTEGA, C. 2021. Genome Characterization of *Lactiplantibacillus plantarum* Strain UTNGt2 Originated from *Theobroma grandiflorum* (White Cacao) of Ecuadorian Amazon: Antimicrobial Peptides from Safety to Potential Applications. **Antibiotics**, v. 10, n.4, abr, 2021, doi: 10.3390/antibiotics10040383
- 11) GONÇALVES DE ALMEIDA, O. G.; VITULO, N.; PEREIRA DE MARTINIS, E. C.; FELIS, G. E. Pangenome analyses of LuxS-coding genes and enzymatic repertoires in cocoa-related lactic acid bacteria. **Genomics**, v. 113, n. 4, jun, 2021, doi.org/10.1016/j.ygeno.2021.04.010
- 12) SAHOO, S.; MAHAPATRA S. R.; MISRA, N.; SUAR, M. 2021. Application of genomics, transcriptomics, and proteomics in probiotic research. In: **Probiotic Beverages**. Elsevier, Amsterdã, Países Baixos, p. 235-256, doi: 10.1016/B978-0-12-818588-9.00009-7
- 13) REBAZA-CARDENAS, T.; SILVA-CAJALEÓN, K.; SABATER, C.; DELGADO, S.; MONTES-VILLANUEVA, N. 2021. Ruas-Madiedo, P. "Masato de Yuca" and "Chicha de Siete Semillas" Two Traditional Vegetable Fermented Beverages from Peru as Source for the Isolation of Potential Probiotic Bacteria. **Probiotics Antimicrob Proteins**. doi: 10.1007/s12602-021-09836-x
- 14) SURACHAT K.; DEACHAMAG P.; KANTACHOTE D.; WONGLAPSUWAN M.; JEENKEAWPIAM K.; CHUKAMNER A. 2021. In silico comparative genomics analysis of *Lactiplantibacillus plantarum* DW12, a potential gamma-aminobutyric acid (GABA)-producing strain. **Microbiological Research**. <https://doi.org/10.1016/j.micres.2021.126833>
- 15) ALMEIDA, E. C. P. DE MARTINIS. 2021. Metagenome-Assembled Genomes Contribute to Unraveling of the Microbiome of Cocoa Fermentation. **Applied and Environmental Microbiology**. doi: 10.1128/AEM.00584-21
- 16) YETIMAN, A.; KESKİN, A.; DARENDELI, B.; KOTIL, S.; ORTAKCI, O., DOGAN, O. 2021. Characterization of genomic, physiological, and probiotic features of *Lactiplantibacillus plantarum* DY46 strain isolated from Turkish fermented turnip juice (Shalgam). **Research Square**. doi: 10.21203/rs.3.rs-854638/v1.
- 17) CARPI, F.; COMAN, M.; SILVI, S.; PICCIOLINI, M.; VERDENELLI, M.; NAPOLIONI, V. 2021. Comprehensive pan-genome analysis of *Lactiplantibacillus plantarum* complete genomes. **Journal of Applied Microbiology**. doi: 10.1111/jam.15199
- 18) KUMAR, S.; KUMARI, N.; TALUKDAR, D.; KOTHIDAR, A.; SARKAR, M.; MEHTA, O., KSHETRAPAL, P.; WADHWA, N.; THIRUVENGADAM, R.; KOUNDINYA DESIRAJU, B.; BALAKRISH NAIR, G.; BHATNAGAR, S.; MUKHERJEE, S.; DAS, B. 2021. The Vaginal Microbial

Signatures of Preterm Birth Delivery in Indian Women. **Frontiers in Cellular and Infection Microbiology**. doi: doi.org/10.3389/fcimb.2021.622474

19) YETIMAN, A.; KESKİN, A.; DARENDELI, B. N.; KOTIL, S. E.; ORTAKCI, F.; DOGON, M. Characterization of genomic, physiological, and probiotic features *Lactiplantibacillus plantarum* DY46 strain isolated from fermented Shalgam beverage. 2021. **Food Bioscience**. doi: 10.1016/j.fbio.2021.101499

A maioria das publicações estão na categoria Ciências biológicas (75%), segundo o Dimensions (<https://badge.dimensions.ai/details/id/pub.1113363198>). De acordo com o mesmo link, a publicação foi citada 4,88 mais vezes que a média de outras da área. Isso revela o relevante impacto científico da publicação.

5. CONSIDERAÇÕES FINAIS

A genômica comparativa revolucionou a forma de analisar os seres vivos, principalmente microrganismos. A quantidade de sequências genômicas públicas é extensa e sua análise contribui de diferentes formas na compreensão da biologia bacteriana, mesmo que a maioria delas sejam sequências rascunho, como àquelas pertencentes à *B. fragilis*. Nossos resultados mostraram que tanto o uso de genomas completos quanto os *drafts* são informativos para o estudo da evolução genômica bacteriana. No entanto, o uso de ferramentas de bioinformática adequadas e atualizadas, com mais acurácia e recursos, é fundamental na exploração de dados genômicos.

BIBLIOGRAFIA

- ABRAHAM, E.P.; CHAIN, E. An enzyme from bacteria able to destroy penicillin. 1940. *Nature*, v. 10, n. 4, p. 677-8, ago, 1988.
- ABRIOUEL, H.; PÉREZ MONTORO, B.; CASADO MUÑOZ, M. et al. In Silico genomic insights into aspects of food safety and defense mechanisms of a potentially probiotic *Lactobacillus Pentosus* MP-10 isolated from brines of naturally fermented Aloreña green table olives. *PLoS One*, v. 12, n. 6, p. e0176801, jun, 2017.
- ADEGBOLA, R.A. Bacterial adhesion and pathogenicity. *Afr J Med Med Sci*, v. 17, n. 2, p. 63-69, jun, 1988.
- AGUIRRE, M.; COLLINS, M.D. Lactic acid bacteria and human clinical infection. *J Appl Bacteriol*, v. 75, n. 2, p. 95-107, ago, 1993.
- AHMED, M.B.; ZHONG, L.; SHEN, C. et al. Colistin and its role in the Era of antibiotic resistance: an extended review (2000-2019). *Emerg Microbes Infec*, v. 9, n. 1, p. 868-885, dez, 2020.
- AHN, K.B.; BAIK, J.B.; PARK, O. et al. *Lactobacillus plantarum* lipoteichoic acid inhibits biofilm formation of *Streptococcus mutans*. *PLoS One*, v. 8, n. 2, e0192694, fev, 2018.
- AKHTER, S.; AZIZ, R.K.; EDWARDS, R.A. PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. *Nucleic Acids Res*, v. 40, n. 16, p. e126, set, 2012.
- ALAN, Y.; TOPALCENGİZ, Z.; DIĞRAK, M. Biogenic amine and fermentation metabolite production assessments of *Lactobacillus plantarum* isolates for naturally fermented pickles. *LWT - Food Science and Technology*, v. 98, p. 322-328, dez, 2018.
- ALCARAZ, L. D. Pan-genomics: Unmasking the gene diversity hidden in the bacteria species. *PeerJ PrePrints*, v. 2, p. e113v2, jan, 2014.
- ALCOCK, B.P.; RAPHENYA A.R.; LAU, T.T. et al. CARD 2020: Antibiotic Resistome Surveillance with the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Research*, v. 48, n. D1, p. D517-D525, out, 2020.
- ALEKSHUN, M.N.; LEVY, S.B. Molecular mechanisms of antibacterial multidrug resistance. *Cell*, v. 128, n. 6, p. 1037-1050, mar, 2007.

- ALIKHAN, N.; PETTY, N.K.; ZAKOUR, N.L.B. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics*, v. 12, p. 402, ago, 2011.
- ALTSCHUL, S.; MADDEN, T. L.; SCHÄFFER, A. A. et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res*, v. 25, n. 17, p. 3389–3402, set, 1997.
- ALVAREZ, C.A.; JONES, M.B.; HAMBOR, J. et al. Characterization of Polysaccharide A Response Reveals Interferon Responsive Gene Signature and Immunomodulatory Marker Expression. *Front Immunol*, v. 11, n. 556813, out, 2020.
- ANDERSSSEN, E.L.; DIEP, D.B.; NES, I.F. et al. Antagonistic activity of *Lactobacillus plantarum* C11: two new two-peptide bacteriocins, plantaricins EF and JK, and the induction factor plantaricin A. *Appl Environ Microb*, v. 64, n. 6, p. 2269–2272, jun, 1998.
- ANKA, N.; SYDENHAMB, T. V.; IVERSENC, L. H. et al. Characterisation of a multidrug-resistant *Bacteroides fragilis* isolate recovered from blood of a patient in Denmark using. *Int J Antimicrob Agents*, v. 46, n. 1, p. 117-120, jul, 2015.
- ANUKAM, K.C.; MACKLAIM, J.M.; GLOOR, G.B. et al. Genome sequence of *Lactobacillus pentosus* KCA1: vaginal isolate from a healthy premenopausal woman. *PLoS One*, v. 8, n. 3, p. E59239, mar, 2013.
- ARBOLEYA, S.; BOTTACINI, F.; O'CONNELL-MOTHERWAY, M. et al. Gene-trait matching across the *Bifidobacterium longum* pan-genome reveals considerable diversity in carbohydrate catabolism among human infant strains. *BMC Genomics*, v. 19, n. 1, p. :33, jan, 2018.
- ARNDT, D.; GRANT, J.; MARCU, A. et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res*, v. 44, W1, p. 16-21, jul, 2016.
- ARNDT, D.; MARCU, A.; LIANG, Y. et al. PHAST, PHASTER and PHASTEST: Tools for finding prophage in bacterial genomes. *Brief Bioinform*, v. 20, n. 4, p. 1560-1567, jul, 2019.
- ATASSI, F.; SERVIN, A.L. Individual and co-operative roles of lactic acid and hydrogen peroxide in the killing activity of enteric strain *Lactobacillus johnsonii* NCC933 and vaginal strain *Lactobacillus gasseri* KS120.1 against enteric, uropathogenic and vaginosis-associated pathogens. *FEMS Microbiol Lett*, v. 304, n. 1, p. 29-38, mar, 2010.
- AW, W.; FUKUDA, S. Toward the comprehensive understanding of the gut ecosystem via metabolomics-based integrated omics approach. *Semin Immunopathol*, v. 37, n. 1, p. 5–16, out, 2015.

- BACIC, M.; PARKER, A. C.; STAGG, J. et al. Genetic and structural analysis of the *Bacteroides* conjugative transposon CTn341. *J Bacteriol*, v. 187, n. 8, p.2858-69, abr, 2005.
- BALCIUNAS, E.M.; CASTILLO MARTINEZ, F.A.; TODOROV, S.D. et al. Novel biotechnological applications of bacteriocins: a review. *Food Control*, v. 32, n. 1, p. 134–142, jul, 2013.
- BAMIAS, G.; COMINELLI, F. Cytokines and intestinal inflammation. *Curr Opin Gastroenterol*, v. 32, n. 6 p. 437-442, nov, 2016.
- BARRANGOU, R.; FREMAUX C.; DEVEAU, H. et al. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, v. 315, n. 5819, p. 1709-1712, mar, 2007.
- BEHERA, S.S.; RAY, R.C.; ZDOLEC, N. *Lactobacillus plantarum* with Functional Properties: An Approach to Increase Safety and Shelf-Life of Fermented Foods. *Biomed Res Int*. PMID: PMC5994577, mar, 2018.
- BENZ, J.; MEINHART, A. Antibacterial effector/immunity systems: it's just the tip of the iceberg. *Curr Opin Microbiol*, v. 17, p. 1-10, fev, 2014.
- BIAGI, E.; NYLUND, L.; CANDELA, M. et al. Through ageing, and beyond: gut microbiota and inflammatory status in seniors and centenarians. *PLoS One*, v. 5, n. 5, e10667, mai, 2010.
- BLAND, C.; RAMSEY, T.L.; SABREE, F. et al. CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics*, v. 8, n. 209, Jun, 2007.
- BOBAY, L.M.; TOUCHON, M.; ROCHA, E.P. Pervasive domestication of defective prophages by bacteria. *Proc Natl Acad Sci U S A*, v. 111, n. 33, p. 12127-12132, ago, 2014.
- BONEV, B.B.; BREUKINK, E.; SWIEZEWSKA, E. et al. Targeting extracellular pyrophosphates underpins the high selectivity of nisin. *FASEB J*, v. 18, n. 15, p. 1862–1869, dez, 2004.
- BONHEYO, G.T.; HUND, B.D.; SHOEMAKER, N.B. et al. Transfer region of a *Bacteroides* conjugative transposon contains regulatory as well as structural genes. *Plasmid*, v. 46, n. 3, p. :202–209, nov, 2001.
- BORTOLAIA, V.; KAAS, R.F.; RUPPE, E. et al. ResFinder 4.0 for predictions of phenotypes from genotypes. *J Antimicrob Chemother*, v. 75, p. : 3491–3500, ago, 2020.
- BOTTA, C., ACQUADRO, A., GREPPI, A. et al. Genomic assessment in *Lactobacillus plantarum* links the butyrogenic pathway with glutamine metabolism. *Sci Rep*, v. 7, n. 15975, nov, 2017.
- BOUTER, K.; BAKKER, G.J.; LEVIN, E. et al. Differential metabolic effects of oral butyrate treatment in lean versus metabolic syndrome subjects. *Clin Transl Gastroenterol*, v. 9, n. 5, p. 155, mai, 2018.

- BOYD, E.F. Bacteriophage-encoded bacterial virulence factors and phage-pathogenicity island interactions. *Adv Virus Res*, v. 82, p. 91-118, dez, 2012.
- BOZDOGAN, B.; APPELBAUM, P.C. Macrolide resistance in streptococci and *Haemophilus influenzae*. *Clinics in Laboratory Medicine*, v. 24, n. 2, p. 455-475, jun, 2004.
- BRANSCOMB, E.; PREDKI, P. On the high value of low standards. *J Bacteriol*, v. 184, n. 23, p. 6406-6409, dez, 2002.
- BURGESS, C.M.; SMID, E.J.; RUTTEN, G.; VAN SINDEREN, D. A general method for selection of riboflavin-overproducing food grade micro-organisms. *Microb Cell Fact*, v. 18, 5, n. 24, jun, 2006.
- BUSH, K. Past and present perspectives on β -lactamases. *Antimicrob Agents Chemother*, v. 24, n. 62: p. 01076-18, Set, 2018.
- CAMACHO, C.C.; OULOIRIS, G.; VAGYAN, V. et al. BLAST plus : architecture and applications. *BMC Bioinformatics*, v. 10, n. 421, dec, 2009.
- CAO, Y.; ROCHA, E. R.; SMITH, C. J. Efficient utilization of complex N-linked glycans is a selective advantage for *Bacteroides fragilis* in extraintestinal infections. *Proc Natl Acad Sci U S A*, v. 111, n. 35, p. 12901-12906, ago, 2014.
- CAPLICE, E.; FITZGERALD, G.F. Food fermentations: role of microorganisms in food production and preservation. *Int J Food Microbiol*, v. 50, n. 1-2, p. 131-149, set, 1999.
- CAPY, P.; GASPERI, G.; BIÉMONT, C. et al. Stress and transposable elements: coevolution of useful parasites? *Heredity*, v. 85:101-106, mar, 2000.
- CARR, V. R.; SHKOPOROV, A.; HILL, C. et al. Probing the Mobilome: Discoveries in the Dynamic Microbiome. *Trends Microbiol*, v. 29, n. 2, p. 158-170, fev, 2021.
- CARVER, T.J.; RUTHERFORD, K.M.; BERRIMAN, M. et al. ACT: the Artemis Comparison Tool. *Bioinformatics*, v. 21, n. 16, ago, 2005.
- CASCALES, E.; BUCHANAN, S. K.; DUCHÉ, D. et al. Colicin biology. *Microbiol Mol Biol Rev*, v. 71, n. 1, p. 158-229, mar, 2007.
- CASTRO-GONZÁLEZ, J.M.; CASTRO, P.; SANDOVAL, H. et al. Probiotic Lactobacilli Precautions. *Front Microbiol*, v. 10, n. 375, mar, 2019.
- CATALOLUK, O.; GOGEBAKAN, B. Presence of drug resistance in intestinal lactobacilli of dairy and human origin in Turkey. *FEMS Microbiol Lett*, v. 236, n. 1, p. 7-12, jul, 2004.
- CHASKES, S; AUSTIN, R. Practical Handbook of Microbiology. Goldman, E.; Lorrence, H. (Org.). 3^a ed. New York, US, CRC Press Taylor & Francis Group. 2015. p. 46.
- CHATZIDAKI-LIVANIS, M. N.; GEVA-ZATORSKY, N.; COMSTOCK, L. E. *Bacteroides fragilis* type VI secretion systems use novel effector and immunity proteins to antagonize human

- gut Bacteroidales species. *Proc Natl Acad Sci U S A*, v. 113, n. 13, p. 3627–3632, mar, 2016.
- CHAUDHARI, N. M.; GUPTA, V. K.; DUTTA, C. BPGA- an ultra-fast pan-genome analysis pipeline. *Sci Rep*, v. 6, n. 24373, abr, 2016.
- CHE, Y.; YANG, Y.; XU, X. et al. Conjugative plasmids interact with insertion sequences to shape the horizontal transfer of antimicrobial resistance genes. *Proc Natl Acad Sci U S A*, v. 118, n. 6, p. e2008731118, fev, 2021.
- CHEN L, YANG J, YU J. et al. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res*, v. 33, p. D325-8, 2005.
- CHEN, L.; ZOU, Y.; PENG, J. et al. *Lactobacillus acidophilus* suppresses colitis-associated activation of the IL-23/Th17 axis. *J Immunol Res*. n. 909514, Abr, 2015.
- CHENG C.W.; LIN H.S.; YE, J.Y. et al. Clinical significance of and outcomes for *Bacteroides fragilis* bacteremia. *J Microbiol Immunol Infect*, v. 42, n. 3, p. 243–250, jun, 2009.
- CHENG, W.T.; KANTILAL, H. K.; DAVAMANI, F. The Mechanism of *Bacteroides fragilis* Toxin Contributes to Colon Cancer Formation. *Malays J Med Sci*, v. 27, n. 4, p. 9-21, Jul, 2020.
- CHIMENTO, D. P.; MOHANTY, A. K.; KADNER, R. J. et al. Substrate-induced transmembrane signaling in the cobalamin transporter Btu. *Nat Struct Biol*, v. 10, n. 5, p. 394-401, mai, 2003.
- CHO, G.S.; HUCH, M.; HANAK, A. et al. Genetic analysis of the plantaricin EFI locus of *Lactobacillus plantarum* PCS20 reveals an unusual plantaricin E gene sequence as a result of mutation. *Int J Food Microbiol*, v. 141(Supl 1), p. S117-24, jul, 2010.
- CHOI, V.M.; HERROU, J.; HECHT A.L. et al. Activation of *Bacteroides fragilis* toxin by a novel bacterial protease contributes to anaerobic sepsis in mice. *Nat Med*, v. 22, n. 5, p. 563-7, mai, 2016.
- CHRISTIE, P.J.; WHITAKER, N.; GONZÁLEZ-RIVERA, C. Mechanism and structure of the bacterial type IV secretion systems. *Biochim Biophys Acta*, v. 1843, n. 1578-91, ago, 2014.
- CLEMENTE, J.C.; URSELL, L.K.; PARFREY, L. et al. The impact of the gut microbiota on human health: an integrative view. *Cell*, v. 148, n. 6, p. 1258–70, mar, 2012.
- COMSTOCK, L.E.; COYNE, M.J.; TZIANABOS, A.O. et al. Interstrain Variation of the Polysaccharide B Biosynthesis Locus of *Bacteroides fragilis*: Characterization of the Region from Strain 638R. *J Bacteriol*, v. 181, n. 19, p. 6192–6196, out, 1999.

- CONNELL, S.R.; TRACZ, D.M.; NIERHAUS, K.H. et al. Ribosomal protection proteins and their mechanism of tetracycline resistance. *Antimicrob Agents Chemother*, v. 47, n. 12, p. 3675-81, dez, 2003.
- CORTÉS-ZAVALA, O.; LÓPEZ-MALO, A.; HERNÁNDEZ-MENDOZA, A. et al. Antifungal activity of lactobacilli and its relationship with 3-phenyllactic acid production. *Int J Food Microbiol*, v. 173, p. 30-35, dez, 2014.
- COTTER, P.D.; HILL, C.; ROSS, R.P. Bacteriocins: developing innate immunity for food. *Nat Rev Microbiol*, v. 3, p. 777-788, out, 2005.
- COUVIN D, BERNHEIM, A.; TOFFANO-NIOCHE, C. et al. CRISPRCasFinder, an update of CRISPRfinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res*, v. 46, n. W1, p. W246-W251, jul, 2018.
- COYNE, M.J.; COMSTOCK, L.E. Type VI Secretion Systems and the Gut Microbiota. *Microbiol Spectr*, v. 7, n. 2, mar, 2019.
- COYNE, M.J.; TZIANABOS, A. O.; MALLORY, B. C. et al. Polysaccharide Biosynthesis Locus Required for Virulence of *Bacteroides fragilis*. *Infect Immun*, v. 69, n. 7, p. 4342-4350, jul, 2001.
- CRANDALL, K.A.; KELSEY, C.R.; IMAMICHI, H. et al. Parallel evolution of drug resistance in HIV: failure of nonsynonymous/synonymous substitution rate ratio to detect selection. *Mol Biol Evol*, v. 16, n. 3, p. 372-382, mar, 1999.
- CROFTS, T.S.; WANG, B.; SPIVAK, A. et al. Shared strategies for β -lactam catabolism in the soil microbiome. *Nat Chem Biol*, v. 14, n. 6, p. 556-564, abr, 2018.
- DAHMUS, J.D.; KOTLER, D.L.; KASTENBERG, D.M. et al. The gut microbiome and colorectal cancer: a review of bacterial pathogenesis. *J Gastrointest Oncol*, v. 9, n. 4, p. 769-77, ago, 2018.
- DARLING, A.E.; MAU, B.; BLATTNER, F.R. et al. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome research*, v. 14, n. 7, p. 1394-1403, jul, 2004.
- DARLING, A.E.; MAU, B.; PERNA, N.T. ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One*, v. 5, n.6, p. e11147, jun, 2010.
- DARRIBA, D.; TABOADA, G. L.; DOALLO, R. et al. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, v. 9, n 8, p. 772, jul, 2012.
- DAVIDSON, A.L.; DASSA, E.; ORELLE, C. et al. Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev*, v. 72, n. 2, p. 317-364, jun, 2008.

- DAVIS, J.J.; WATTAM, A.R.; AZIZ, R.K. The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Res*, v. 48, n. D1, p. D606–D612, jan, 2020.
- DE ARAUZ, L.J.; JOZALA, A.F.; MAZSZOLA, P.G, et al. Nisin biotechnological production and application: a review. *Trends Food Sci Tech*, v. 20, n. 3, p. 146–154, abr, 2009.
- DE MAN, T.J.B.; STAJICH, J.E.; KUBICEK, C.P. et al. Small genome of the fungus *Escovopsis weberi*, a specialized disease agent of ant agriculture. *Proc Natl Acad Sci U S A*, v. 113, n. 13, p. 3567-72, mar, 2016.
- DE VUYST, L.; VANDAMME, E.J. Bacteriocins of lactic acid bacteria microbiology, genetics and applications. *Microbiology, Genetics and Applications*. London, UK: Blackie Academic & Professional, Londres, Reino Unido, 1994. 194-199 p.
- DELVES-BROUGHTON, J.; BLACKBURN, P.; EVANS, R.J. et al. Applications of the bacteriocin, nisin. *Antonie Van Leeuwenhoek*, v. 69, p. 193–202, fev, 1996.
- DEPLANCKE, B.; GASKINS, H.R. Microbial modulation of innate defense: goblet cells and the intestinal mucus layer. *Am J Clin Nutr*, v. 73, n. 6, p. 1131S-1141S, jun, 2001.
- DIEP, D.B.; HÅVARSTEIN, L.S.; NES, I.F. Characterization of the locus responsible for the bacteriocin production in *Lactobacillus plantarum* C11. *J Bacteriol*, v. 178, n. 15, p. 4472-83, ago, 1996.
- DIEP, D.B.; STRAUME, D.; KJOS, M. et al. An overview of the mosaic bacteriocin *pln* loci from *Lactobacillus plantarum*. *Peptides*, v. 30, n. 8, p. 1562-74, mai, 2009.
- DING, W.; BAUMDICKER, F.; NEHER, R.A. PanX: pan-genome analysis and exploration. *Nucleic Acids Res*, v. 46, n. 1, e5, Jan, 2018.
- DIRIX, G.; MONSIEURS, P.; DOMBRECHT, B. et al. Peptide signal molecules and bacteriocins in Gram-negative bacteria: a genome-wide in silico screening for peptides containing a double-glycine leader sequence and their cognate transporters. *Peptides*, v. 25, n. 9, p. 1425-1440, set, 2004.
- DOMINGUES, S.; HARMS, K.; FRICKE, W.F. et al. Natural Transformation Facilitates Transfer of Transposons, Integrons and Gene Cassettes between Bacterial Species. *PLoS Pathog*, v. 8, n. 8, e1002837, ago, 2012.
- DOMINGUEZ-BELLO, M.G.; COSTELLO, E.K.; CONTRERAS, M. et al. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A*, v. 107, n. 26, p. 11971-11975, jun, 2010.
- DRAGO, L.; RODIGHIERO, V.; MATTINA, R. et al. In vitro selection and transferability of antibiotic resistance in the probiotic strain *Lactobacillus reuteri* DSM 17938. *J Chemother*, v. 23, n. 6, p. 371–373, dez, 2011.

- DRIDER, D.; FIMLAND, G.; HECHARD, Y. et al. The continuing story of class IIa bacteriocins. *Microbiol Mol Biol Rev*, v. 70, n. 2, p. 564–82, jun, 2006.
- ECKBURG, P.B.; BIK, E.M.; BERNSTEIN, C.N. et al. Diversity of the human intestinal microbial flora. *Science*, v. 308, n. 5728, p. 1635–38, mar, 2005.
- EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*, v. 32, v. 5, p. 1792–7, mar, 2004.
- EDGAR, R.C. PILER-CR: Fast and accurate identification of CRISPR repeats. *BMC Bioinformatics*, v. 8, 18, jan, 2007.
- EDGAR, R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, v. 26, n. 19, p. 2460–2461, out, 2010.
- EDLIN, G.; LIN, L.; BITNER, R. Reproductive fitness of P1, P2, and Mu lysogens of *Escherichia coli*. *J Virol*, v. 21, n. 2, p. 560–564, fev, 1977.
- ELMER, G.W.; MCFARLAND, L.V.; SURAWICZ, C.M. et al. Behavior of *Saccharomyces boulardii* in recurrent *Clostridium difficile* disease patients. *Aliment Pharmacol Ther*, v. 13, n. 12, p. 1663–1668, dez, 1999.
- ELSAGHIR, H.; REDDIVARI, A. K. R. *Bacteroides Fragilis*. 2020. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK553032/>
- EMSER, K.; BARBOSA, J.; TEIXEIRA, P. et al. *Lactobacillus plantarum* survival during the osmotic dehydration and storage of probiotic cut apple. *J Funct Foods*, v. 38, p. 519–528, nov, 2017.
- ERACLIO, G.; RICCI, G.; FORTINA, M.G. Insertion sequence elements in *Lactococcus garvieae*. *Gene*, v. 555, n. 2, p. 291–296, jan, 2015.
- EXCOFFIER, L.; LISCHER, E.L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour*, v. 10, n. 3, p. 564–567, mai, 2010.
- FALAGAS, M.E.; KASIAKOU, S.K. COLISTIN: the revival of polymyxins for the management of multidrug-resistant gram-negative bacterial infections. *Clin Infect Dis*, v. 40, n. 9, p. 1333–1341, mai, 2005.
- FALLANI, M.; AMARRI, S.; UUSIJARVI, A. et al. Determinants of the human infant intestinal microbiota after the introduction of first complementary foods in infant samples from five European centres. *Microbiology*, v. 157 (Pt 5), p. 1385–1392, may, 2011.
- FELIS, G.E.; DELLAGLIO, F. Taxonomy of lactobacilli and bifidobacteria. *Curr Issues Intest Microbiol*, v. 8, n. 2, p. 44–61, set, 2007.

- FINEGOLD, S.M. Overview of clinically important anaerobes. *Clinical Infectious Diseases*, v. 20, no. 2, pp. 205-207, jul, 1995.
- FOURIE, K.R.; WILSON, H.L. Understanding GroEL and DnaK Stress Response Proteins as Antigens for Bacterial Diseases. *Vaccines*, v. 8, n.4, p. 773, 2020.
- FOX, P.F. 1993. Cheese: An overview. In: *Cheese: Chemistry, Physics and Microbiology*. Fox, P.F. (Ed.). Segunda Edição, Chapman and Hall, Londres, Reino Unido. 36 p.
- FRANCKE C.; KORMELINK T. G.; HAGEMEIJER, Y. et al. (2011) Comparative analyses imply that the enigmatic sigma factor 54 is a central controller of the bacterial exterior. *BMC Genomics* v. 12, p. 385.
- FRANCO, A.A. The *Bacteroides fragilis* pathogenicity island is contained in a putative novel conjugative transposon. *J Bacteriol*, v. 186, n. 18, p. 6077-92, set, 2004.
- FRANCO, A.A.; CHENG, R.K.; GOODMAN, A. et al. Modulation of bft expression by the *Bacteroides fragilis* pathogenicity island and its flanking region. *Mol Microbiol*, v. 45, n.4, p. 1067-1077, ago, 2002.
- FRAS, P.; CAMPOS, F.M.; HOGG, T. et al. Production of volatile phenols by *Lactobacillus plantarum* in wine conditions. *Biotechnol Lett*, v. 36, n. 2, p. 281-285, fev, 2014.
- FURUKAWA, K. Oxygenases and dehalogenases: Molecular approaches to efficient degradation of chlorinated environmental pollutants. *Biosci Biotechnol Bioch*, v. 70, n. 10, p. 2335-2348, out, 2006.
- GÄNZLE, M.G. Lactic metabolism revisited: metabolism of lactic acid bacteria in food fermentations and food spoilage. *Curr Opin Food Sci*, v. 2, p. 106-117, abr, 2015.
- GARCÍA-BAYONA, L.; COYNE, M.J.; COMSTOCK, L.E. Mobile Type VI secretion system loci of the gut Bacteroidales display extensive intra-ecosystem transfer, multi-species spread and geographical clustering. *PLoS Genet*, v. 17, n. 4, p. e1009541, abr, 2021.
- GARCÍA, N.; GUTIÉRREZ, G.; LORENZO, M. et al. Gene Context and DNA rearrangements in the carbapenemase locus of division II strains of *Bacteroides fragilis*. *Antimicrob Agents Chemother*, v. 53, n. 6, p. 2677-2678, jul, 2009.
- GARRITY, G.M.; BELL, J.A.; LILBURN, T.M. 2004. Taxonomic outline of the prokaryotes. In *Bergey's Manual of Systematic Bacteriology*, 2º Ed. Disponível em: <https://laurieximenez.files.wordpress.com/2012/02/3-taxonomia-de-prokaryotes-bergeys-msb-2004.pdf>. Acesso em: 14 mar. 2021.
- ASIUNAS, G.; BARRANGOU, R.; HORVATH, P. et al. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A*, v. 109, n. 39, p. E2579-E2586, set, 2012.

- GASKINS, H.R. Immunological aspects of host/microbiota interactions at the intestinal epithelium. In: Mackie RI, White BA, Isaacson RE. *Gastrointestinal Microbiology*, v. 2, p. 537–87, 1997.
- GAYNES, R. The Discovery of Penicillin—New Insights After More Than 75 Years of Clinical Use. *Emerg Infect Dis*, v. 23, n. 5, p. 849–853, mai, 2017.
- GILLESPIE, J.H. 1991. Moderate Selection. In: *The causes of molecular evolution*. Oxford University Press (Ed.), Nova Iorque. 299 p.
- GÖDEKE, J.; PAUL, K.; LASSAK, J. et al. Phage-induced lysis enhances biofilm formation in *Shewanella oneidensis* MR-1. *ISME J*, v. 5, n. 4, p. 613–626, abr, 2011.
- GOLDSTEIN, E.J. Anaerobic bacteremia. *Clin Infect Dis*, v. 23, p. S97–S101, dec, 1996.
- GOLDSTEIN, E.J.; TYRRELL, K.L.; CITRON, D.M. *Lactobacillus* species: taxonomic complexity and controversial susceptibilities. *Clin Infect Dis*, v. 60 Supl 2, p. S98–107, mai, 2015.
- GRISSA, I.; VERGNAUD, G.; POURCEL, C. CRISPRFinder : a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res*, v. 35, p. W52–57, jul, 2007.
- GUARNER, F.; SCHAAFSMA, G.J. Probiotics. *Int J Food Microbiol*, v. 39, n. 3, 237–238, fev, 1998.
- GUÉDON, G.; LIBANTE, V.; COLUZZI, C. et al. The Obscure World of Integrative and Mobilizable Elements, Highly Widespread Elements that Pirate Bacterial Conjugative Systems. *Genes (Basel)*, v. 8, n. 11, p. 337, nov, 2017.
- GUINDON, S.; DUFAYARD, J. F.; LEFORT V. et al. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology*, v. 59, n. 3, p. 307–321, mai, 2010.
- GUPTA, S.K.; PADMANABHAN, B.R.; DIENE, S.M. et al. Arg-annot, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother*, v. 58, n. 1, p. 212–220, out, 2014.
- GUTACKER, M.; VALSANGIACOMO, C.; PIFFARETTI, J.C. Identification of two genetic groups in *Bacteroides fragilis* by multilocus enzyme electrophoresis: distribution of antibiotic resistance (*cfiA*, *cepA*) and enterotoxin (*bft*) encoding genes. *Microbiology*, v. 146, p. 1241–1254, mai, 2000.
- HADFIELD, J.; CROUCHER, N.J.; GOATER, R.J. et al. Phandango: an interactive viewer for bacterial population genomics. *Bioinformatics*, v. 34, n. 2, p. 292–293, jan, 2018.
- HAGGOUD, A.; REYSSET, G.; AZEDDOUG, H. et al. Nucleotide sequence analysis of two 5-nitroimidazole resistance determinants from *Bacteroides* strains and of a new

- insertion sequence upstream of the two genes. *Antimicrob Agents Chemother*, v. 38, n. 5, p. 1047-51, mai, 1994.
- HAGHI, F.; GOLI, E.; MIRZAEI, B. et al. The association between fecal enterotoxigenic *B. fragilis* with colorectal cancer. *BMC Cancer*, v.19, n. 1, p. 879, set, 2019.
- HAGHIGHAT, L.; CRUM-CIANFLON, N.F. The potential risks of probiotics among HIV-infected persons: bacteremia due to *Lactobacillus acidophilus* and review of the literature. *Int J STD AIDS*, v. 27, n. 13, p. 1223–1230, nov, 2016.
- HARDY, R.G.; MELTZER, S.J.; JANKOWSKI, J.A. ABC of colorectal cancer. Molecular basis for risk factors. *BMJ*, v. 321, n. 7265, p. 886 –889, out, 2000.
- HARVEY, K.L.; JAROCKI, V.M.; CHARLES, I.G. et al. The Diverse Functional Roles of Elongation Factor Tu (EF-Tu) in Microbial Pathogenesis. *Front Microbiol*, v. 10, p. 2351, nov, 2019.
- HASAN, K.Z.; PATHELA, P.; ALAM, K. et al. Aetiology of diarrhoea in a birth cohort of children aged 0–2 year(s) in rural Mirzapur, Bangladesh. *J Health Popul Nutr*, v. 24, n. 1, 24, p. 25–35, mai, 2006.
- HERROU, J.; CHOI, V.M.; BUBECK WARDENBURG, J. et al. Activation mechanism of the *Bacteroides fragilis* cysteine peptidase, fragipain. *Biochemistry*, v. 55, n. 29, p. 4077–4084, jun, 2016.
- HO, B.T.; DONG, T.G.; MEKALANOS, J. J. A view to a kill: the bacterial type VI secretion system. *Cell Host Microbe*, v. 15, n. 1, p. 9–21, jan, 2014.
- HOLDEN, M.T.; FEIL, E.J.; LINDSAY, J.A. et al. Complete genomes of two clinical *Staphylococcus aureus* strains: Evidence for the rapid evolution of virulence and drug resistance. *Proc Natl Acad Sci U S A*, v. 101, n. 26, 9786-9791, jun, 2004.
- HORVATH, P.; BARRANGOU, R. CRISPR/Cas, the immune system of bacteria and archaea. *Science*, v. 327, n. 5962, p. 167–70, jan, 2010.
- HU, Y.; YANG, X.; QIN, J. et al. Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat Commun*. n. 4, 2151, jul, 2013.
- HUBERT, J.C.; KAMMERER, B. Determination of chromosome size and number of *rrn* loci in *Lactobacillus plantarum* by pulsed-field gel electrophoresis. *FEMS Microbiol Lett*, v. 120, n-1-2, p. 51-56, jul, 1994.
- HUNT, V.L.; TSAI, I.J.; COGHLAN, A. et al. The genomic basis of parasitism in the *Strongyloides* clade of nematodes. *Nat Genet*, v. 48, n. 3, p. 299–307, mar, 2016.
- JACOBSEN, L.; WILCKS, A.; HAMMER, K. et al. Horizontal transfer of *tet*(M) and *erm*(B) resistance plasmids from food strains of *Lactobacillus plantarum* to *Enterococcus*

- faecalis JH2-2 in the gastrointestinal tract of gnotobiotic rats. *FEMS Microbiol Ecol*, v.59, n.1, p. 158-66, jan, 2007.
- JEANNOT, K.; BOLARD, A.; PLÉSIAT, P. Resistance to polymyxins in Gram-negative organisms. *Int J Antimicrob Agents*, v. 49, n. 5, p. 526-535, mai, 2017.
- JIA, B.; RAPHENYA, A. R.; ALCOCK, B. et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res*, v. 45, n. D1, p. D566-D573, out, 2017.
- JIA, F.F.; ZHANG, L.J.; PANG, X.H. 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*, v. 109, n. 5-6, p. 432-437, out, 2017.
- JIA, W.; LI, H.; ZHAO, L. et al. Gut microbiota: a potential new territory for drug targeting. *Nat Rev Drug Discov*, v. 7, n. 2, p. 123-9, fev, 2008.
- JIANG, J.; YANG, B.; ROSS, R.P. et al. Comparative Genomics of *Pediococcus pentosaceus* Isolated From Different Niches Reveals Genetic Diversity in Carbohydrate Metabolism and Immune System. *Front Microbiol*, v. 11, p. 253, nov, 2020.
- JIMENEZ-DIAZ, R.; RUIZ-BARBA, J.L.; CATHCART, D.P. et al. Purification and partial amino acid sequence of plantaricin S, a bacteriocin produced by *Lactobacillus plantarum* LPCO10, the activity of which depends on the complementary action of two peptides. *Appl Environ Microb*, v. 61, n. 12, p. 4459-4463, dez, 1995.
- JOHANSSON, M.H. K.; BORTOLAIA, V.; TANSIRICHAIIYA, S. et al. Detection of mobile genetic elements associated with antibiotic resistance in *Salmonella enterica* using a newly developed web tool: MobileElementFinder. *J Antimicrob Chemother*, v. 76, n. 1, p. 101-109, jan, 2021.
- JOHNSON, C. M.; GROSSMAN, A. D. Integrative and Conjugative Elements (ICEs): What They Do and How They Work. *Annu Rev Genet*, v. 49, p. 577-601, out, 2015.
- JOHNSON, E.L.; HEAVER, S.L.; WALTERS, W.A. et al. Microbiome and metabolic disease: revisiting the bacterial phylum Bacteroidetes. *J Mol Med*, v. 95, n. 1, p. 1-8, jan, 2017.
- JOZALA, A.F.; LENCASTRE NOVAES, L.C.; PESSOA JUNIOR, A. 2015. Nisin, Concepts, Compounds and the Alternatives of Antibacterials, Varaprasad Bobbarala, IntechOpen, Disponível em: <https://www.intechopen.com/books/concepts-compounds-and-the-alternatives-of-antibacterials/nisin>. Acesso em: 14 mar. 2021.

- KAKIKAWA, M.; OKI, M.; TADOKORO H. et al. Cloning and nucleotide sequence of the major capsid proteins of Lactobacillus bacteriophage phi gle. *Gene*, v. 175, n. 1-2, p. 157-165, out, 1996.
- KALITA, D.; SAIKIA, S.; GAUTAM, G. et al. Characteristics of synbiotic spray dried powder of litchi juice with Lactobacillus plantarum and different carrier materials. *LWT- Food Science and Technology*. LWT, v. 87, p. 351–360, jan, 2018.
- KASTNER, S.; PERRETEN, V.; BLEULER, H. et al. Antibiotic susceptibility patterns and resistance genes of starter cultures and probiotic bacteria used in food. *Syst Appl Microbiol*, v. 29, n. 2, p. 145-155, mar, 2006.
- KATO, N.; LIU, C. X.; KATO, H. et al. A new subtype of the metalloprotease toxin gene and the incidence of the three bft subtypes among Bacteroides fragilis isolates in Japan. *FEMS Microbiology letters*, v. 182, n. 1, p. 171-6, jan, 2000.
- KAYSER, F.H.; BIENZ, K.A.; ECKERT, J. et al. 2006. General Bacteriology. In: Medical Microbiology. Stuttgart (Ed.), Nova Iorque. pp.152-156.
- KEMPF, M.; ROLAIN, J.M. Emergence of resistance to carbapenems in Acinetobacter baumannii in Europe: clinical impact and therapeutic options. *Int J Antimicrob Agents*, v. 39, n. 2, p. 105–14, fev, 2011.
- KHALAF, H.; NAKKA, S.S.; SANDÉN, C. et al. Antibacterial effects of Lactobacillus and bacteriocin PLNC8 $\alpha\beta$ on the periodontal pathogen Porphyromonas gingivalis. *BMC Microbiol*, v. 16, n. 1, p. 188, ago, 2016.
- KHAZAIE, K.; ZADEH, M.; KHAN, M.W. et al. Abating colon cancer polyposis by Lactobacillus acidophilus deficient in lipoteichoic acid. *Proc Natl Acad Sci U S A*, v. 109, n. 26, p. 10462–10467, jun, 2012.
- KLAENHAMMER, T.R. Genetics of bacteriocins produced by lactic acid bacteria. *FEMS Microbiol Rev*, v. 12, n. 1-3, p. 39-86, set, 1993.
- KLEEREBEZEM, M.; BOEKHORST, J.; VAN KRANENBURG, R. et al. Complete genome sequence of Lactobacillus plantarum WCFS1. *Proc Natl Acad Sci U S A*, v. 100, n. 4, p. 1990-1995, fev, 2003.
- KLEIN, G.; PACK, A.; BONAPARTE, C. et al. Taxonomy and physiology of probiotic lactic acid bacteria. *Int J Food Microbiol*, v. 41, n. 2, p. 103-25, mai, 1998.
- KO, S.H.; RHO, D.J.; JEON, J.I. et al. Bacteroides fragilis Enterotoxin Upregulates Heme Oxygenase-1 in Intestinal Epithelial Cells via a Mitogen-Activated Protein Kinase- and NF-kappaB-Dependent Pathway, Leading to Modulation of Apoptosis. *Infect Immun*, v. 84, n. 9, p. 2541–2554, ago, 2016.

- KODAIRA, K.I.; OKI, M.; KAKIKAWA, M. et al. Genome structure of the *Lactobacillus* temperate phage phi g1e: the whole genome sequence and the putative promoter/repressor system. *Gene*, v. 187, n. 1, p. 45-53, mar, 1997.
- KONSTANTINIDIS, K.T.; TIEDJE, J.M. Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci U S A*, v. 102, n. 7, p. 2567-2572, fev, 2005.
- KORBER, B.; FISCHER, W.M.; GNANAKARAN, S. et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell*, v. 182, n. 4, p. 812-827, ago, 2020.
- KOROPATKIN, N.M.; CAMERON, E.A.; MARTENS, E.C. How glycan metabolism shapes the human gut microbiota. *Nat Rev Microbiol*, v. 10, n. 5, p. 323-335, abr, 2012.
- KOSAKOVSKY, POND, S.L.; FROST, S.D.W. Not So Different After All: A Comparison of Methods for Detecting Amino Acid Sites Under Selection. *Mol Biol Evol*, v. 22, n. 5, p. 1208-1222, mai, 2005.
- KOSAKOVSKY, POND, S.L.; FRONT, S.D.W. MUSE, S.V. et al.. HyPhy: hypothesis testing using phylogenies. *Bioinformatics*, v. 21, n. 5, p. 676-679, mar, 2005.
- KOSAKOVSKY POND, S.L.; POON A.F.Y. VELAZQUEZ, R. et al. HyPhy 2.5- A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Mol Biol Evol*, v. 37, n. 1, p. 295-299, jan, 2020.
- KOSAKOVSKY, POND, S.L.; POSADA, D.; GRAVENOR, M.B. et al. GARD: a genetic algorithm for recombination detection. *Bioinformatics*, v. 22, n. 24, p. 3096-3098, dez, 2006.
- KOSAKOVSKY, POND, S.L.; POSADA, D.; GRAVENOR, M.B. et al. Automated Phylogenetic Detection of Recombination Using a Genetic Algorithm. *Mol Biol Evol*, v. 23, n. 10, p. 1891-1901, out, 2006.
- KUMAR, S.; STECHER, G.; LI, M.; KNYAZ, C. et al. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol Biol Evol*, v. 35, n. 6, p. 1547-1549, jun, 2018.
- KURIHARA, T.; ESAKI, N. Bacterial hydrolytic dehalogenases and related enzymes: occurrences, reaction mechanisms, and applications. *Chem Rec*, v. 8, n. 2, p. 67-74, mar, 2008.
- KWON, E.K.; KANG, G.D.; KIM W.K. et al. *Lactobacillus plantarum* LC27 and *Bifidobacterium longum* LC67 simultaneously alleviate ethanol-induced gastritis and hepatic injury in mice. *J Funct Foods*, v. 38, p. 389-398, jul, 2019.
- LABRIE, S.J.; SAMSON, J.; MOINEAU, S. Bacteriophage resistance mechanisms. *Nat Rev Microbiol*, v. 8, n. 5, p. 317-327, mai, 2010.

- LAY, C.; DORÉ, J.; RIGOTTIER-GOIS, L. Separation of bacteria of the *Clostridium leptum* subgroup from the human colonic microbiota by fluorescence-activated cell sorting or group specific PCR using 16S rRNA gene oligonucleotides. *FEMS Microbiol Ecol*, v. 60, n. 513-520, abr, 2007.
- LEE, G.; CHAKRABORTY, U.; GEBHART, D. et al. F-Type Bacteriocins of *Listeria monocytogenes*: a New Class of Phage Tail-Like Structures Reveals Broad Parallel Coevolution between Tailed Bacteriophages and High-Molecular-Weight Bacteriocins. *J Bacteriol*, v. 198, n.20, p. 2784-2793, out, 2016.
- LÉPINE, G.; LACROIX, J. M.; WALKER, C. B. A Progulske-Fox. Sequencing of a tet(Q) gene isolated from *Bacteroides fragilis* 1126. *Antimicrob Agents Chemother*, v. 37, n. 9, p. 2037-2041, set, 1993.
- LEROY, F.; DE VUYST, L. Lactic acid bacteria as functional starter cultures for the food fermentation industry. *Trends Food Sci Technol*, v. 15, n. 2, p. 67-78, fev, 2004.
- LEY, R.E.; PETERSON, D.A.; GORDON, J.I. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell*, v. 124, n. 4, p. 837-48, fev, 2006.
- LI, L.; STOECKERT, C.J.; ROOS, D.S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res*, v. 13, n. 9, p. 2178-89, set, 2003.
- LI, W; WU, C. I; LUO, C.C.A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol Biol Evol*, v. 2, n. 2, p. 150-174, mar, 1985.
- LIN, C.F.; FUNG, Z.F.; WU, C.L. et al. Molecular characterization of a plasmid-borne (pTC82) chloramphenicol resistance determinant (cat-TC) from *Lactobacillus reuteri* G4. *Plasmid*, v. 36, n. 2, p. 116-24, set, 1996.
- LIU, B.; ZHENG, D.D.; JIN, Q. et al. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res*, v. 47, n. D1, p. D687-D692, jan, 2019.
- LIU, M.; LI, X.; XIE. et al. ICEberg 2.0: an updated database of bacterial integrative and conjugative elements. *Nucleic Acids Res*, v. 47, n. D1, p. D660-D665, jan, 2019.
- LJUNGH, A.; WADSTRÖM, T. Lactic acid bacteria as probiotics. *Curr Issues Intest Microbiol*, v. 7, n. 2, p. 73-89, set, 2006.
- LUBELSKI, J.; RINK, R.; KHUSAINOV, R. et al. Biosynthesis, immunity, regulation, mode of action and engineering of the model lantibiotic nisin. *Cell Mol Life Sci*, v. 65, n. 3, p. 455-76, fev, 2008.
- LUONG, M.L.; SAREYYUPOGLU, B.; NGUYEN M.H. et al. *Lactobacillus* probiotic use in cardiothoracic transplant recipients: a link to invasive *Lactobacillus* infection? *Transplant Infect Dis*, v. 12, n. 6, p. 561-564, dez, 2010.

- MA, J.; LI, Z.; ZHANG, W. et al. Comparison of gut microbiota in exclusively breast-fed and formula-fed babies: a study of 91 term infants. *Sci Rep*, v. 10, n. 15792, set, 2020.
- MACFARLANE, G.T.; MACFARLANE, S. Fermentation in the human large intestine: its physiologic consequences and the potential contribution of prebiotics. *J Clin Gastroenterol*, v. 45, p. 120–7, nov, 2011.
- MAKAROVA, K.S.; HAFT, D.H.; BARRANGOU, R. et al. Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol*, v. 9, n. 6, p. 467-77, jun, 2011.
- MALDONADO, A.; RUIZ-BARBA, J.L.; JIMENEZ-DIAZ, R. Purification and genetic characterization of plantaricin NC8, a novel coculture-inducible two-peptide bacteriocin from *Lactobacillus plantarum* NC8. *Appl Environ Microbiol*. 69:383-389, jan, 2003.
- MÁRMOL, I.; SÁNCHEZ-DE-DIEGO, C.; PRADILLA DIESTE, A. et al. Colorectal Carcinoma: A General Overview and Future Perspectives in Colorectal Cancer. *Int J Mol Sci*, v. 18, n. 1, p. 197, Jan, 2017.
- MARTINO, M.E.; BAYJANOV, J.R.; CAFFREY, B.E. et al. Nomadic lifestyle of *Lactobacillus plantarum* revealed by comparative genomics of 54 strains isolated from different habitats. *Environ Microbiol*, v. 18, n. 12, p. 4974-4989, dez, 2016.
- MESSIER, W.; STEWART, C.B. Episodic adaptive evolution of primate lysozymes. *Nature*, v. 385, n. 6612, p. 151-4, jan. 1997.
- MASUDA, M.; IDE, M.; UTSUMI, H. et al. Production potency of folate, vitamin B(12), and thiamine by lactic acid bacteria isolated from Japanese pickles. *Biosci Biotechnol Biochem*, v. 76, n. 11, p. 2061-7, nov, 2012.
- MAZMANIAN, S.K.; LIU, C.H.; TZIANABOS, A.O. et al. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell*, v. 122, n. 1, p. 107-18, jul, 2005.
- MAZMANIAN, S.K.; ROUND, J. L.; KASPER, D. L. A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature*, v. 453, n. 7195, p. 620-5, mai, 2008.
- MEDINI, D.; DONATI, C.; TETTELIN, H. et al. The microbial pan-genome. *Curr Opin Genet Dev*, v. 15, n. 6, p. 589-94, dez, 2005.
- MEINI, S.; LAUREANO, R.; FANI, L. et al. Breakthrough *Lactobacillus rhamnosus* GG bacteremia associated with probiotic use in an adult patient with severe active ulcerative colitis: case report and review of the literature. *Infection*, v. 43, n. 6, p. 777–781, dec, 2015.
- MIRA, A.; MARTÍN-CUADRADO, A.B.; D'AURIA, G. et al. The bacterial pan-genome: a new paradigm in microbiology. *Int Microbiol*, v. 13, n. 2, p. 45-57, jan, 2010.

- MOAZED, D.; NOLLER, H.F. Chloramphenicol, erythromycin, carbomycin and vernamycin B protect overlapping sites in the peptidyl transferase region of 23S ribosomal RNA. *Biochimie*, v. 69, n. 8, p. 879-84, ago, 1987.
- MOLENAAR, D.; BRINGEL, F.; SCHUREN, F.H. et al. Exploring *Lactobacillus plantarum* genome diversity by using microarrays. *J Bacteriol*, v. 187, n. 17, p. 6119-6127, set, 2005.
- MONCRIEF, J.S.; DUNCAN, A.J.; WRIGHT, R.L. et al. Molecular characterization of the fragilysin pathogenicity islet of enterotoxigenic *Bacteroides fragilis*. *Infect Immun*, v. 66, n. 4, p. 1735-1739, abr, 1998.
- MONCRIEF, J.S.; OBISO, R.J.R.; BARROSO, L.A. et al. The enterotoxin of *Bacteroides fragilis* is a metalloprotease. *Infect Immun*, v. 63, n. 1, p. 175-81, jan, 1995.
- MORENO-ARRIBAS, M.V.; POLO, M. C.; JORGANES, F. et al. Screening of biogenic amine production by lactic acid bacteria isolated from grape must and wine. *Int. J. Food Microbiol.* vol. 84, n. 1, p. 117-123, jun, 2003.
- MOZZI, F. 2016. Lactic Acid Bacteria. *Encyclopedia of Food and Health*. Elsevier (Ed.), Amsterdã, Países Baixos, 501 p.
- MUNOZ-LOPEZ, M.; GARCIA-PEREZ, J.L. DNA transposons: nature and applications in genomics. *Curr Genomics*, v. 11, n. 2, p. 115-128, abr, 2010.
- MUNOZ-PRICE, L.S.; POIREL, L.; BONOMO, R.A. et al. Clinical epidemiology of the global expansion of *Klebsiella pneumoniae* carbapenemases. *Lancet Infect Dis*, v. 13, n. 9, p. 785-96, set, 2013.
- MUÑOZ-PROVENCIO, D.; PÉREZ-MARTÍNEZ, G.; MONEDERO, V. Characterization of a fibronectin-binding protein from *Lactobacillus casei* BL23. *J Appl Microbiol.* 108, n. 3, p. 1050-1059, mar, 2010.
- MURRELL, B.; MOOLA, S.; MABONA, A. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol*, v. 30, n. 5, p. 1196-205, mai, 2013.
- MURRELL, B.; WEAVER, S.; SMITH, M.D. et al. Gene-Wide Identification of Episodic Selection. *Mol Biol Evol*, v. 32, n. 5, p. 1365-1371, mai, 2015.
- MURRELL, B.; WERTHEIM, J.O.; MOOLA, S. et al. Detecting Individual Sites Subject to Episodic Diversifying Selection. *PLoS Genetics*, v. 8, no. 7, p. e1002764, jul, 2012.
- MYERS L.L.; SHOOP, D.S. Association of enterotoxigenic *Bacteroides fragilis* with diarrheal disease in young pigs. *Am. J. Vet. Res*, v. 48, n. 5, p. 774-775, mai, 1987.

- MYERS, L.L.; FIREHAMMER, B.D.; SHOOP, D.S. et al. *Bacteroides fragilis*: a possible cause of acute diarrheal disease in newborn lambs. *Infect Immun*, v. 44, n. 2, p. 241-244, mai, 1984.
- MYERS, L.L.; SHOOP, D.S.; FIREHAMMER, B.D. et al. Association of enterotoxigenic *Bacteroides fragilis* with diarrheal disease in calves. *J Infect Dis*, v. 152, n. 6; p. 1344-1347, dez, 1985.
- MYERS, L.L.; SHOOP, D.S.; STACKHOUSE, L.L. Isolation of enterotoxigenic *Bacteroides fragilis* from humans with diarrhea. *J Clin Microbiol*, v. 25, n. 2, p. 2330-3, dez, 1987.
- NAKAMURA, I.; AOKI, K.; MIURA, Y. et al. Fatal sepsis caused by multidrug-resistant *Bacteroides fragilis*, harboring a *cfiA* gene and an upstream insertion sequence element, in Japan. *Anaerobe*, v. 44, p. 36-39, abr, 2017.
- NAMI, Y.; ABDULLAH, N.; HAGHSHEENAS, B. et al. Assessment of probiotic potential and anticancer activity of newly isolated vaginal bacterium *Lactobacillus plantarum* 5BL. *Microbiol Immunol*, v. 58, n. 9, p. 492-502, set, 2014.
- NEAL-MCKINNEY, J.M.; LU, X.; DUONG, T. et al. Production of organic acids by probiotic lactobacilli can be used to reduce pathogen load in poultry. *PLoS ONE*, v. 7, n. 9, e43928, set, 2012.
- NGUYEN, T.V.; VAN, P.L.; HUYNH, C.L. et al. Diarrhea caused by enterotoxigenic *Bacteroides fragilis* in children less than 5 years of age in Hanoi, Vietnam. *Anaerobe*, p. 11, n. 1-2, p. 109-114, fev, 2005.
- NIELSEN, R.; HELLMANN, I.; HUBISZ, M. et al. Recent and ongoing selection in the human genome. *Nat Rev Genet*, v. 8, n. 11, p. 857-868, 2007.
- NIELSEN, R.; YANG, Z. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics*, v. 148, n. 3, p. 929-936, mar, 1998.
- NIESTĘPSKI, S.; HARNISZ, M.; KORZENIEWSKA, E. et al. *Bacteroides* spp. - clinical significance, antibiotic resistance and identification methods. *Postępy Mikrobiol*, v. 56, n. 1, p. 67-76, abr, 2017.
- NISSEN-MEYER, J.; ROGNE, P.; OPPEGARD, C. et al. Structure- function relationships of the non-lanthionine-containing peptide (class II) bacteriocins produced by Gram-positive bacteria. *Curr Pharm Biotechnol*, v. 10, n. 1, p. 19-37, jan, 2009.
- ONDERDONK, A.B.; KASPER, D.L.; CISNEROS, R.L. et al. The capsular polysaccharide of *Bacteroides fragilis* as a virulence factor: comparison of the pathogenic potential of encapsulated and unencapsulated strains. *J Infect Dis*, v. 136, n. 1, p. 82-89, jul, 1977.

- OOI, L.G.; LIONG, M.T. Cholesterol-lowering effects of probiotics and prebiotics: a review of in vivo and in vitro findings. *Int J Mol Sci*, v. 11, n. 6, p. 2499-2522, jun, 2010.
- OTTOMAN, N.; SMIDT, H.; DE VOS, W.M. et al. The function of our microbiota: who is out there and what do they do? *Front Cell Infect Microbiol*, v. 9, n. 2, p. 104, ago, 2012.
- PARIDA, S.; WU, S.; SIDDHARTH, S. et al. A pro-carcinogenic colon microbe promotes breast tumorigenesis and metastatic progression and concomitantly activates Notch and β catenin axes. *Cancer Discov*. CD-20-0537, jan, 2021.
- PARK, E.K.; KIM, H.; JIN, H.J. Phylogenetic analysis of rRNA methyltransferases, Erm and KsgA, as related to antibiotic resistance. *FEMS Microbiology Letters*, v. 309, n. 2, p. 151-162, ago, 2010.
- PARK, J.S.; CHOI, J.W.; JHUN, J. et al. *Lactobacillus acidophilus* Improves Intestinal Inflammation in an Acute Colitis Mouse Model by Regulation of Th17 and Treg Cell Balance and Fibrosis Development. *J Med Food*, v. 21, n. 3, p. 215-224, mar, 2018.
- PARTE, A.C.; SARDÀ CARBASSE, J.; MEIER-KOLTHOFF, J.P. et al. List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol*. v. 70, n. 11, p. 5607-5612, nov, 2020.
- PARTRIDGE, S.R.; TSAFNAT, G.; COIERA, E. et al. Gene cassettes and cassette arrays in mobile resistance integrons. *FEMS Microbiol Rev*, v. 33, n. 4, p. 757-784, abr, 2009.
- PATHELA, P.; HASAN, K.Z.; ROY, E. et al. Enterotoxigenic *Bacteroides fragilis*-associated diarrhea in children 0-2 years of age in rural Bangladesh. *J Infect Dis*, v. 191, n. 8, p. 1245-52, abr, 2005.
- PEARSON, W. R. An introduction to sequence similarity ("homology") searching. *Curr Protoc Bioinform*, v. 3, cap. 3.1, jun, 2013.
- PEREZ-MUÑOZ, M.E.; ARRIETA, M.C.; RAMER-TAIT, A.E. et al. A critical assessment of the "sterile womb" and "in utero colonization" hypotheses: implications for research on the pioneer infant microbiome. *Microbiome*, v. 5, n.1, apr, 2017.
- PFEILER, E.A.; KLAENHAMMER, T.R. The genomics of lactic acid bacteria. *TRENDS in Microbiology*, v. 15, n. 12, p. 546-553, dez, 2007.
- PIDDOCK, L.J. Multidrug-resistance efflux pumps - not just for resistance. *Nat Rev Microbiol*, v. 4, p. 629-36, ago, 2006.
- PIERCE, J.V.; BERNSTEIN, H.D. Genomic Diversity of Enterotoxigenic Strains of *Bacteroides fragilis*. *PLoS One*, v. 11, n. 6, p. e0158171, jun, 016.
- PIMENTA, A.L.; RACHER, K.; JAMIESON, L. et al. Mutations in HlyD, part of the type 1 translocator for hemolysin secretion, affect the folding of the secreted toxin. *J. Bacteriol*, v. 187, n. 21, p. 7471-80, nov, 2005.

- POMPEI, A.; CORDISCO, L.; AMARETTI, A. et al. Folate production by bifidobacteria as a potential probiotic property. *Appl Environ Microbiol*, v. 73, n. 1, p. 179–85, jan, 2007.
- PONTAROTTI, P.; RAOULT, D. Why do arthropods secrete β -lactams? *Int J Antimicrob Agents*, v. 53, n. 4, p. 370, nov, 2019.
- RAY, R.C.; JOSHI V.K. Fermented Foods: Past, present and future scenario. In: *Microorganisms and Fermentation of Traditional Foods*. Ray R. C., Montet D. (Ed.). CRC Press; Boca Raton, EUA, 2014. 2 p.
- READ, T.D.; USSERY, D.W. Opening the pan-genomics box. *Curr Op Microbiol*, v. 9, p. 496–498, dez, 2006.
- REEVES, A.R.; D'ELIA, J.N.; FRIAS, J. et al. A *Bacteroides thetaiotaomicron* outer membrane protein that is essential for utilization of maltooligosaccharides and starch. *J Bacteriol*, v. 178, n. 3, p. 823–30, fev, 1996.
- RICCI, A.; ALLENDE, A.; BOLTON, D. et al. Update of the list of QPS-recommended biological agents intentionally added to food or feed as notified to EFSA 5: suitability of taxonomic units notified to EFSA until September 2016. *EFSA Journal*, v. 15, n. 3, set, 2016.
- ROACH, M.J.; MCNAIR, K.; GILES, S.K et al. Philympics 2021: Prophage Predictions Perplex Programs. *F1000Research*, v. 10, n. 758, 2021.
- RODRÍGUEZ, J. M.; MARTÍNEZ, M.I.; KOK, J. Pediocin PA-1, a wide-spectrum bacteriocin from lactic acid bacteria. *Crit Rev Food Sci Nutr*, v. 42, n. 2, p. 91–121, mar, 2002.
- ROGERS, M.B.; BENNETT, T.K.; PAYNE, C.M. et al. Insertional activation of *cepA* leads to high-level beta-lactamase expression in *Bacteroides fragilis* clinical isolates. *J Bacteriol*, v. 176, n. 14, p. 4376–4384, jul, 1994.
- ROSEN, B.P. Biochemistry of arsenic detoxification. *FEBS Lett*, v. 529, n. 1, p. 86–92, out, 2002.
- ROUND, J.L.; MAZMANIAN, S.K. Inducible Foxp3⁺ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proc Natl Acad Sci U S A*, v. 107, n. 27, p. 12204–12209, jul, 2010.
- RUSSELL, A.B.; HOOD, R.D.; BUI, N.K. et al. Type VI secretion delivers bacteriolytic effectors to target cells. *Nature*, v. 475, n. 7356, p. 343–347, jul, 2011.
- RUSSELL, A.B.; PETERSON, S.B.; MOUGOUS, J.D. Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol*, v. 12, n. 2, p. 137–48, fev, 2014.

- SACK, R.B.; ALBERT, M.J.; ALAM K. et al. Isolation of enterotoxigenic *Bacteroides fragilis* from Bangladeshi children with diarrhea: a controlled study. *J Clin Microbiol*, v. 32, n. 4, p. 960–3, abr, 1994.
- SAELIM, K.; JAMPAPHAENG, K.; MANEERAT, S. Functional properties of *Lactobacillus plantarum* S0/7 isolated fermented stinky bean (Sa Taw Dong) and its use as a starter culture. *J Funct Foods*, v. 38, p. 370–377, nov, 2017.
- SAHL, H.; BIERBAUM, G. Lantibiotics: Biosynthesis and biological activities of uniquely modified peptides from Gram-positive bacteria. *Ann Rev Microbiol*, v. 52: 41-79, out, 1998.
- SALVETTI, E.; HARRIS, H.M.B.; FELIS, G.E. et al. Comparative Genomics of the Genus *Lactobacillus* Reveals Robust Phylogroups That Provide the Basis for Reclassification. *Appl Environ Microbiol*, v. 84, n. 20, out, 2018.
- SALVETTI, E.; TORRIANI, S.; FELIS, G.E. The genus *Lactobacillus*: a taxonomic update. *Probiotics Antimicrob Proteins*, v. 4, n. 4, p. 217–226, dez, 2012.
- SALYERS, A.A. *Bacteroides* of the human lower intestinal tract. *Annu Rev Microbiol*, v. 38, p. 293–313, out, 1984.
- SAN JOAQUIN, V.H.; GRIFFIS, J.C.; LEE, C. et al. Association of *Bacteroides fragilis* with childhood diarrhea. *Scand J Infect Dis*, v. 27, n. 3, p. 211-5, dez, 1995.
- SAYERS, S.; LI, L.; ONG, E. et al. Victors: a web-based knowledge base of virulence factors in human and animal pathogens. *Nucleic Acid Research*, v. 47, n. D1, p. D693-D700, jan, 2019.
- SCHLEIFER, K.H.; STACKEBRANDT, E. Molecular systematics of prokaryotes. *Annu Rev Microbiol*, v. 37, p. 143–187, out, 1983.
- SCHNEIDER, D.; LENSKI, R.E. Dynamics of insertion sequence elements during experimental evolution of bacteria. *Res. Microbiol*, v. 155, n. 5, p. 319-327, jun, 2004.
- SCIENCE, U.o.M.S.o.M.I.f.G. (2013). GSCID Genomic Sequencing Center for Infectious Disease [Online]. NCBI. Disponível em: http://gscid.igs.umaryland.edu/wp.php?wp=comparative_genomics_of_human_mucosal_isolates_of_the_enterotoxigenic_bacteroides_fragilis_group. Acesso em: 14 de setembro de 2021.
- SEARS, C.L. The toxins of *Bacteroides fragilis*. *Toxicon*, v. 39, n. 11, p. 1737-46, nov, 2001.
- SEARS, C.L.; FRANCO, A.; WU, S. *Bacteroides fragilis* toxins. In: The Comprehensive Sourcebook of Bacterial Protein Toxins. Alouf J, Popoff M, Ed. 3rd ed. Academic Press; Oxford, Reino Unido, 2005. p. 535-546.

- SEARS, C.L.; GEIS, A.L.; HOUSSEAU, F. *Bacteroides fragilis* subverts mucosal biology: from symbiont to colon carcinogenesis. *J Clin Invest*, v. 124, n. 10, out, 2014.
- SEARS, C.L.; MYERS, L.L.; LAZENBY, A et al. Enterotoxigenic *Bacteroides fragilis*: a Rogue among Symbiotes. *Clin Infect Dis. sup.* p. S142-148, jun, 1995.
- SEDDIK, H.A.; BENDALI, F.; GANCEL, F. et al. *Lactobacillus plantarum* and Its Probiotic and Food Potentialities. *Probiotics Antimicrob Proteins*, v. 9, n. 2, p. 111-122, jun, 2017.
- SEEMANN, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, v. 30, n. 14, p. 2068–2069, jul, 2014.
- SENESI, S.; CELANDRONI, F.; TAVANTI, A. et al. Molecular characterization and identification of *Bacillus clausii* strains marketed for use in oral bacteriotherapy. *Appl Environ Microbiol*, v. 67, n. 2, p. 834–839, fev, 2001.
- SHEN, G.; WU, J.; YE, B. et al. Gut Microbiota-Derived Metabolites in the Development of Diseases. *Can J Infect Dis Med Microbiol*. PMID: PMC7815404, Jan, 2021.
- SHIN, J.M.; GWAK, J.W.; KAMARAJAN, P. et al. Biomedical applications of nisin. *J Appl Microbiol*, v. 120, n. 6, p. 1449-1465, jun, 2016.
- SHIRYAEV, S.A.; ALESHIN A.E.; MURANAKA, N. et al. Structural and functional diversity of metalloproteinases encoded by the *Bacteroides fragilis* pathogenicity island. *FEBS J*, v. 281, n. 11, p. 2487-502, jun, 2014.
- SHMAKOV, S.; ABUDAYYEH, O.O.; MAKAROVA, K.S. et al. Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. *Mol Cell*, v. 60, n. 3, p. 385–397, nov, 2015.
- SHOKRYAZDAN, P.; SIEO, C.C.; KALAVATHY, R. et al. Probiotic potential of *Lactobacillus* strains with antimicrobial activity against some human pathogenic strains. *Biomed Res Int*. 927268, jul, 2014.
- SIEZEN, R.J.; TZENEVA, V.A.; CASTIONI, A. et al. Phenotypic and genomic diversity of *Lactobacillus plantarum* strains isolated from various environmental niches. *Environ Microbiol*, v. 12, n. 3, p. 758-773, mar, 2010.
- SIEZEN, R.J.; VAN HYLCKAMA VLIEG, J.E.T. Genomic diversity and versatility of *Lactobacillus plantarum*, a natural metabolic engineer. *Microb Cell Fact*, v. 10, n. S3, 2011.
- SIMS, D.; SUDBERY, I.; ILOTT, N.E. et al. Sequencing depth and coverage: key considerations in genomic analyses. *Nature Reviews Genetics*.v. 15, n. 2, p. 121-132, jan, 2014.
- SINGH, V.K.; UTAIDA, S.; JACKSON, L.S. et al. Role for *dnaK* locus in tolerance of multiple stresses in *Staphylococcus aureus*. *Microbiology*, v. 153, n. 9, p. 3162–3173, set, 2017.

- SÓKI, J.; HEDBERG, M.; PATRICK, S. et al. The isolates noted as virulent and multidrug-resistant (VIR/MDR) include an international cluster of MDR *B. fragilis* isolates from five countries. *J Antimicrob Chemother*, v. 71, n. 9, p. 2441–2448, set, 2016.
- SONG, L.; PAN, J.; YANG, Y. et al. Contact-independent killing mediated by a T6SS effector with intrinsic cell-entry properties. *Nat Commun*, v. 423, jan, 2021.
- STINSON, L.F.; BOYCE, M.C.; PAYNE, M.S. et al. The Not-so-Sterile Womb: Evidence That the Human Fetus Is Exposed to Bacteria Prior to Birth. *Front Microbiol*, v. 10, p. 1124, jun, 2019.
- SUN, J.; ZHANG, H.; LIU, Y.H. et al. Towards understanding MCR-like colistin resistance. *Trends Microbiol*, v. 26, n. 9, p. 794–808, set, 2018.
- SUN, Z.; ZHONG, J.; LIANG. et al. Novel mechanism for nisin resistance via proteolytic degradation of nisin by the nisin resistance protein NSR. *Antimicrob Agents Chemother*, v. 53, n. 5, p. 1964–1973, mai, 2009.
- TABASCO, R.; SÁNCHEZ-PATÁN, F.; MONAGAS, M. et al. Effect of grape polyphenols on lactic acid bacteria and bifidobacteria growth: resistance and metabolism. *Food Microbiol*, v. 28, n. 7, p. 1345–1352, out, 2011.
- TAKAHASHI, S.E.; MENDES, A.A.; SALDANHA, E.S.P.B. et al. Efficiency of prebiotics and probiotics on the performance, yield, meat quality and presence of *Salmonella* spp in carcasses of free-range broiler chickens. *Rev Bras Cienc Avic*, v. 7, p. 151–157, n.3, set, 2005.
- TAN, T.Y.; NG L.S.; KWANG, L.L. et al. Clinical characteristics and antimicrobial susceptibilities of anaerobic bacteremia in an acute care hospital. *Anaerobe*, v. 43, p. 69–74, fev, 2017.
- TANG, H.; LYONS, E.; PEDERSEN, B. et al. Screening synteny blocks in pairwise genome comparisons through integer programming. *BMC Bioinformatics*, v. 12, n. 102, abr, 2011.
- TANNOCK, G.W. Probiotics and prebiotics. Where are we going? *Emerg Infect Dis*, v. 9, n. 5, p. 610, mai, 2003.
- TANNOCK, G.W.; LUCHANSKY, J.B.; MILLER, L. et al. Molecular characterization of a plasmid-borne (pGT633) erythromycin resistance determinant (ermGT) from *Lactobacillus reuteri* 100-63. *Plasmid*, v. 31, n. 1, p. 60–71, jan, 1994.
- TETTELIN, H.; MASIGNANI, V.; CIESLEWICZ, M.J. et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". *Proc Natl Acad Sci U S A*, v. 102, n. 39, p. 13950–13955, set, 2005.

- TODOROV, S.D. Bacteriocins from *Lactobacillus plantarum* production, genetic organization and mode of action. *Braz J Microbiol*, v. 40, n. 2, p. 209-221, jun, 2009.
- TOMOYASU, T.; TABATA, A.; IMAKI, H. et al. Role of *Streptococcus intermedius* DnaK chaperone system in stress tolerance and pathogenicity. *Cell Stress Chaperon*, v. 17, n. 1, p. 41-55, jan, 2011.
- TORRIANI, S.; CLEMENTI, F.; VANCANNEYT, M. et al. Differentiation of *Lactobacillus plantarum*, *L. pentosus* and *L. paraplantarum* species by RAPD-PCR and AFLP. *Syst Appl Microbiol*, v. 24, n. 4, p. :554-560, dez, 2001.
- TURNER, D.L.; BRENNAN, L.; MEYER, H.E. et al. Solution structure of plantaricin C, a novel lantibiotic. *Eur J Biochem*, v. 264, n. 3, p. 833-839, set, 1999.
- TZIANABOS, A.O.; CHANDRAKER, A.; KALKA-MOLL, W. et al. Bacterial pathogens induce abscess formation by CD4(+) T-cell activation via the CD28-B7-2 costimulatory pathway. *Infect Immun*, v. 68, n. 12, p. 6650-6655, dez, 2000.
- TZOUVELEKIS, L.S.; MARKOGIANNAKIS, A.; PSICHOGIOU, M. et al. Carbapenemases in *Klebsiella pneumoniae* and other Enterobacteriaceae: an evolving crisis of global dimensions. *Clin Microbiol Rev*, v. 25, n. 4, p. 682-707, out, 2012.
- VACANTE, M.; CIUNI, R.; BASILE F. et al. Gut Microbiota and Colorectal Cancer Development: A Closer Look to the Adenoma-Carcinoma Sequence. *Biomedicines*, v. 8, n. 11, p. 489, nov, 2020.
- VAHABNEZHAD, E.; MOCHON, A.B.; WOZNIAK L.J. et al. *Lactobacillus* bacteremia associated with probiotic use in a pediatric patient with ulcerative colitis. *J Clin Gastroenterol*, v. 47, n. 5, p. 437-9, mai, 2013.
- VALGUARNERA, E.; WARDENBURG, J.B. Good gone bad: One toxin away from. disease for *Bacteroides fragilis*. *J Mol Biol*, v. 432, n. 4, p. 765-785, dec, 2020.
- VAN DEN NIEUWBOER, M.; VAN HEMERT, S.; CLAASSEN, E. et al. *Lactobacillus plantarum* WCFS1 and its host interaction: a dozen years after the genome. *Microb Biotechnol*, v. 9, n. 4, p. 452-465, jul, 2016.
- VAN KRANENBURG, R.; GOLIC, N.; BONGERS, R. et al. Functional analysis of three plasmids from *Lactobacillus plantarum*. *Appl Environ Microb*, v. 71, n. 3, 1223-1230, mar, 2005.
- VANDAMME, P.; POT, B.; GILLIS, M. et al. Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol Rev*, v. 60, n. 2, p. 407-38, jun, 1996.
- VERGARA, I.A.; CHEN, N. Large synteny blocks revealed between *Caenorhabditis Elegans* and *Caenorhabditis Briggsae* genomes using OrthoCluster. *BMC Genomics*, v. 11, n. 516, set, 2010.

- VERSTER, A.J.; ROSS, B.D.; RADEY, M.C. et al. The Landscape of Type VI Secretion across Human Gut Microbiomes Reveals Its Role in Community Composition. *Cell Host & Microbe*, v. 22, n. 3, p. 411-419.e4, set, 2017.
- VITTI, J.J.; GROSSMAN, S.R.; SABETI, P.C. Detecting natural selection in genomic data. *Annu Rev Genet*, v. 47: p. 97–120, nov, 2013.
- VRIEZE, A.; VAN NOOD, E.; HOLLEMAN, F. et al. Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology*, v. 143, n. 4, p. 913-6, out, 2012.
- WANG, J.; JI, H.; ZHANG, D. et al. Assessment of probiotic properties of *Lactobacillus plantarum* ZLP001 isolated from gastrointestinal tract of weaning pigs. *Afr J Biotechnol*, v. 10, n. 54, p. 11303-11308, set, 2011.
- WANG, X.; KIM, Y.; MA, Q. et al. Cryptic prophages help bacteria cope with adverse environments. *Nat Commun*, v. 1, n. 9, p. 147, dez, 2010.
- WANG, Y.; CAO, X.; FENG, Y. et al. Environment-induced conformational and functional changes of l-2-haloacid dehalogenase. *J Biosci Bioeng*, v. 121:n. 5, p. 491-496, mai, 2016.
- WANG, Y.; FENG, Y.; CAO. et al. Insights into the molecular mechanism of dehalogenation catalyzed by D-2-haloacid dehalogenase from crystal structures. *Sci Rep*, v. 8, n. 1, 1454, jan, 2018.
- WAREHAM, D.W.; WILKS, M.; AHMED, D. et al. Anaerobic sepsis due to multidrug-resistant *Bacteroides fragilis*: microbiological cure and clinical response with linezolid therapy. *Clin Infect Dis*, v. 40, n. 7, p. 67–68, abr, 2005.
- WATTAM, A.R.; DAVIS, J.J.; ASSAF, R. et al. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Res*, v. 45, n. D1, p. D535-D542, nov, 2017.
- WEXLER, H. M. *Bacteroides*: the good, the bad, and the nitty-gritty. *Clin Microbiol Rev*, v. 20, n. 4, p. 593-621, out, 2007.
- WHITAKER, W.B.; RICHARDS, G.P.; BOYD, E.F. Loss of sigma factor RpoN increases intestinal colonization of *Vibrio parahaemolyticus* in an adult mouse model. *Infect Immun*, v. 82, n. 2, p. 544-556, fev, 2014.
- WILLIAMS, T.; KELLEY, C. Gnuplot 4.6.6. An interactive plotting program, March 2018, 2014, Available from: <http://www.gnuplot.info>
- WOLLOWSKI, I.; RECHKEMMER, G.; POOL-ZOBEL, B.L. Protective role of probiotics and prebiotics in colon cancer. *Am J Clin Nutr*, v. 73, p. 451S-455S, fev, 2001.

- WU, S.; LIM, K. C.; HUANG, J. et al. *Bacteroides fragilis* enterotoxin cleaves the zonula adherens protein, E-cadherin. *Proc Natl Acad Sci U S A*, v. 95, n. 25, p. 14979-84, dez, 1998.
- WU, S.; MORIN, P.J.; MAOUYO, D. et al. *Bacteroides fragilis* enterotoxin induces c-Myc expression and cellular proliferation. *Gastroenterology*, v. 124, n. 2, p. 392-400, fev, 2003.
- XU, B.; YANG, Z. PAMLX: a graphical user interface for PAML. *Mol Biol Evol*, v. 30, n. 12, p. 2723-4, dez, 2013.
- XU, J.; MAHOWALD, M.A.; LEY, R.E. et al. Evolution of symbiotic bacteria in the distal human intestine. *PLoS Biol*, v. 5, e156, jul, 2007.
- YADAV, A.K.; TYAGI, A.; KAUSHIK, J.K. Et al. Role of surface layer collagen binding protein from indigenous *Lactobacillus plantarum* 91 in adhesion and its anti-adhesion potential against gut pathogen. *Microbiol Res*, v. 168, n. 10, p. 639-645, dez, 2013.
- YAMADA, K.; HIROTA, M.; NIIMI, Y. et al. Nucleotide sequences and organization of the genes for carotovoricin (Ctv) from *Erwinia carotovora* indicate that Ctv evolved from the same ancestor as *Salmonella typhi* prophage. *Biosci Biotechnol Biochem*, v. 70: n. 9, p. 2236-2247, set, 2006.
- YANG, E.J.; CHANG, H.C. Purification of a new antifungal compound produced by *Lactobacillus plantarum* AF1 isolated from kimchi. *Int J Food Microbiol*, v. 139, n.1-2, p. 139:56-63, abr, 2010.
- YANG, Z. Maximum likelihood estimation on large phylogenies and analysis of adaptive evolution in human influenza virus A. *J Mol Evol*, v. 51, n. 5, p. 423-432, nov, 2000.
- YANG, Z.; NIELSEN, R. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *J Mol Evol*, v. 46, p. 409-418, abr, 1998.
- YEKANI, M.; BAGHI, H.B.; NAGHILI, B. et al. To resist and persist: Important factors in the pathogenesis of *Bacteroides fragilis*. *Microb Pathog*, v. 149, n. 104506, set, 2020.
- YOON, B.H.; JANG, S.H.; CHANG, H. Sequence analysis of the *Lactobacillus* temperate phage Sha1. *Arch Virol*, v. 156, n. 9, p. 1681-1684, sep, 2011.
- YOSHINO, Y.; KITAZAWA, T.; IKEDA, M. et al. Clinical features of *Bacteroides* bacteremia and their association with colorectal carcinoma. *Infection*.v. 40, n. 1, p. 63-67, fev, 2012.
- ZANKARI, E.; HASMAN, H.; COSENTINO, S. et al. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother*, v. 67, n. 11, p. 2640-2644, nov, 2012.

- ZHANG, G.; SVENUNGSSON, B.; KÄRNELL, A. et al. Prevalence of enterotoxigenic *Bacteroides fragilis* in adult patients with diarrhea and healthy controls. *Clin Infect Dis*, v. 29, n. 3, p. 590–594, set, 1999.
- ZHANG, H.; LIU, L.; HAO, Y. et al. Isolation and partial characterization of a bacteriocin produced by *Lactobacillus plantarum* BM-1 isolated from a traditionally fermented Chinese meat product. *Microbiol Immunol*, v. 57, n. 11, 746–755, nov, 2013.
- ZHANG, T.; ZHANG, X.; YE, L. Plasmid Metagenome Reveals High Levels of Antibiotic Resistance Genes and Mobile Genetic Elements in Activated Sludge. *PLoS One*, v. 6, n. 10, 6:e26041, out, 2011.
- ZHAO, J.; TIAN, F.; YAN, S. et al. *Lactobacillus plantarum* CCFM10 alleviating oxidative stress and restoring the gut microbiota in d-galactose-induced aging mice. *Food & Functions*, v. 9, n. 2, p. 917–924, dez, 2018.
- ZHAO, Y.; WU, J.; YANG, J. et al. PGAP: pan-genomes analysis pipeline, *Bioinformatics*, v. 28, n. 3, p. 416–418, fev, 2012.
- ZHENG, J.; RUAN, L.; SUN, M. et al. A genomic view of lactobacilli and pediococci demonstrates that phylogeny matches ecology and physiology. *Appl Environ Microbiol*, v. 81, n. 20, p. 7233–7243, out, 2015.
- ZHENG, J.; WITTOUCK, S.; SALVETTI, E. et al. A taxonomic note on the genus *Lactobacillus*: Description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. *Int J Syst Evol Microbiol*, v. 70, n. 4, p. 2782–2858, abr, 2020.
- ZOU Y.; XUE, W.; LUO, G. et al. 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nat Biotechnol*, v. 37, n. 2, p. 179–185, fev, 2019.

APÉNDICE

Table S1: *B. fragilis* strains used in this study.

Strain	Source	Status	Country	Collection date	Accession
143	Blood	Scaffold	Ecuador	Jan-2017	NZ_SRKB000000000.1
1284	Gut microbiota	Scaffold		2015	NZ_PDCX000000000.1
12905	Gut microbiota	Scaffold	USA	2015	NZ_PDCW000000000.1
078320-1	Faeces	Scaffold	USA	Jan-2015	NZ_PDCV000000000.1
1001175B_160314_C 3	Faeces	Scaffold	USA	Mar-2016	NZ_JADMXXK000000000 .1
1001175st1_C3	Faeces	Scaffold	USA	Mar-2016	NZ_SPGY000000000.1
1001285H_161024_D 4	Faeces	Scaffold	USA	Oct-2016	NZ_JADNGU00000000 00.1
1007-1-F #10 ^a	Faeces	Contig	USA	-	NZ_JGEA000000000.1
1007-1-F #3 ^a	Tissue	Contig	USA	Dec-2007	NZ_JGEB000000000.1
1007-1-F #4 ^a	Tissue	Contig	USA	Dec-2007	NZ_JGDK000000000.1
1007-1-F #5 ^a	Tissue	Contig	USA	Dec-2007	NZ_JGVD000000000.1
1007-1-F #6 ^a	Tissue	Contig	USA	Dec-2007	NZ_JGVF000000000.1
1007-1-F #7 ^a	Faeces	Contig	USA	-	NZ_JGDL000000000.1
1007-1-F #8 ^a	Faeces	Contig	USA	-	NZ_JGCM000000000.1
1007-1-F #9 ^a	Faeces	Contig	USA	-	NZ_JGEC000000000.1
1009-4-F #10 ^a	Faeces	Contig	USA	-	NZ_JGED000000000.1
1009-4-F #7 ^a	Faeces	Contig	USA	-	NZ_JGEE000000000.1
14-106904-1	Appendix	Scaffold	Sweden	2014	NZ_MKXL000000000.2
2_1_16	Gastrointestinal tract	Scaffold	USA	-	NZ_ACPP000000000.1
2_1_56FAA	Gastrointestinal tract	Scaffold	USA	-	NZ_ACWI000000000.1
2-078382-3 ^a	Faeces	Contig	USA	1989	NZ_LIDV000000000.1
2-F-2# 4 ^a	Tissue	Contig	USA	Dec-2011	NZ_JGDM000000000.1
2-F-2# 5 ^a	Tissue	Contig	USA	Dec-2011	NZ_JGCN000000000.1
2-F-2# 7 ^a	Tissue	Contig	USA	Dec-2011	NZ_JGCO000000000.1
20656-2-1 ^a	Faeces	Contig	USA	1987	NZ_LIDU000000000.1

20793-3 ^a	Faeces	Contig	USA	1987	NZ_JGEF00000000.1
3-F-2 #6	Faeces	Contig	USA	Dec-2011	NZ_JGDT00000000.1
320_BFRA	Fluid	Contig	USA	-	NZ_JVLR00000000.1
321_BFRA	Fluid	Scaffold	USA	-	NZ_JVLQ00000000.1
322_BFRA	Fluid	Contig	USA	-	NZ_JVLP00000000.1
3397 N2 ^a	Tissue	Contig	USA	2010	NZ_JGDN00000000.1
3397 N3 ^a	Tissue	Contig	USA	-	NZ_JGEG00000000.1
3397 T10 ^a	Tissue	Contig	USA	-	JGDN00000000
3397 T14 ^a	Tissue	Contig	USA	-	NZ_JGDO00000000.1
34-F-2 #13 ^a	Tissue	Contig	USA	Dec-2012	NZ_JGCQ00000000.1
3719 A10 ^a	Tissue	Contig	USA	Jul-2010	NZ_JGDP00000000.1
3719 T6 ^a	Tissue	Contig	USA	Jul-2010	NZ_JGEH00000000.1
3725 D9 ii ^a	Tissue	Contig	USA	Aug-2010	JGEI00000000
3725 D9(v) ^a	Tissue	Contig	USA	Aug-2010	NZ_JGDQ00000000.1
3774 T13 ^a	Tissue	Contig	USA	Mai-2011	NZ_JGCR00000000.1
3783N1-2 ^a	Tissue	Contig	USA	Jul-2011	NZ_JGCS00000000.1
3783N1-6 ^a	Tissue	Contig	USA	Jul-2011	NZ_JGEU00000000.1
3783N1-8 ^a	R colon	Contig	USA	Jul-2011	NZ_JGDR00000000.1
3783N2-1 ^a	Tissue	Contig	USA	Jul-2011	NZ_JGCT00000000.1
3976T7 ^a	Tissue	Contig	USA	Nov-2011	NZ_JGCU00000000.1
3976T8 ^a	Tissue	Contig	USA	Nov-2011	NZ_JGDS00000000.1
3986 N(B)19 ^a	Tissue	Contig	USA	Feb-2012	NZ_JGCW00000000.1
3986 N(B)22 ^a	Tissue	Contig	USA	Feb-2012	NZ_JGEI00000000.1
3986 N3 ^a	Tissue	Contig	USA	Feb-2012	NZ_JGVG00000000.1
3986 T(B)13 ^a	Tissue	Contig	USA	Feb-2012	NZ_JGEJ00000000.1
3986 T(B)9 ^a	Tissue	Contig	USA	Feb-2012	NZ_JGCX00000000.1
3986T(B)10 ^a	Tissue	Contig	USA	Feb-2012	NZ_JGCV00000000.1
3988 T1 ^a	Tissue	Contig	USA	Mar-2012	NZ_JGCZ00000000.1
3988T(B)14 ^a	Tissue	Contig	USA	Mar-2012	NZ_JGCV00000000.1
3996 N(B) 6 ^a	Tissue	Contig	USA	May-2012	NZ_JGDA00000000.1
3998 T(B) 4 ^a	Tissue	Contig	USA	May-2012	NZ_JGDC00000000.1
3998T(B)3 ^a	Tissue	Contig	USA	2012	NZ_JGDB00000000.1

638R	Abdominal abscess	Complete	UK	-	NC_016776.1
86-5443-2-2 ^a	Faeces	Contig	USA	1997	NZ_LIDS00000000.1
885_BFRA	ICU ^b	Contig	USA	-	NZ_JUPJ00000000.1
894_BFRA	ICU ^b	Contig	USA	-	NZ_JUOZ00000000.1
8E3_BL_hyb ^a	Blood	Contig	USA	-	NZ_CAEUHN00000000.1
915_BFRA	Faeces	Contig	USA	-	NZ_JUOA00000000.1
A7 (UDC12-2) ^a	Tissue	Contig	Thailand	-	NZ_JGEK00000000.1
AD126T_1B	Transverse colon	Contig	USA	2018	NZ_VOHZ00000000.1
AD126T_2B	Purulent sample	Contig	USA	2018	NZ_VOIJ00000000.1
AD135F_1B	Purulent sample	Contig	USA	2018	NZ_VOHV00000000.1
AD135F_2B	Faeces	Contig	USA	2018	NZ_WCHZ00000000.1
AD135F_3B	Purulent sample	Contig	USA	2018	NZ_VOHT00000000.1
AF14-14AC	Purulent sample	Scaffold	China	Oct-2013	NZ_QRZO00000000.1
AF14-26	Purulent sample	Scaffold	China	Oct-2013	NZ_QRZH00000000.1
AF15-10LB	Faeces	Scaffold	China	Nov-2013	NZ_QRYS00000000.1
AF26-6	Purulent sample	Scaffold	China	Dec-2013	NZ_QRTT00000000.1
AF27-10	Purulent sample	Scaffold	China	Dec-2013	NZ_QRTQ00000000.1
AF32-10	Purulent sample	Scaffold	China	Jan-2014	NZ_QRQH00000000.1
am_0171	Purulent sample	Contig	USA	Jan-2015	NZ_RCXN00000000.1
AM13-18	Faeces of an elephant	Scaffold	China	Aug-2013	NZ_QRLH00000000.1
AM15-16	Clinical isolate	Scaffold	China	Sep-2013	NZ_QRKP00000000.1
AM16-16A	Faeces	Scaffold	China	Sep-2013	NZ_QRKL00000000.1
AM17-19	Faeces	Scaffold	China	Sep-2013	NZ_QRJX00000000.1
AM18-6	Faeces	Scaffold	China	Sep-2013	NZ_QRJE00000000.1
AM26-13LB	Faeces	Scaffold	China	Nov-2013	NZ_QSLE00000000.1
AM31-13AC	Faeces	Scaffold	China	Dec-2013	NZ_QSJF00000000.1
AM40-4AC	Faeces	Scaffold	China	Jan-2014	NZ_QSGL00000000.1
AM47-7	Faeces	Scaffold	China	Apr-2014	U30316
ATCC 25285	Apendix abscess	Scaffold	USA	-	NZ_MTGH00000000.1

B1 (UDC16-1) ^a	Tissue	Contig	USA	-	PDCU01000000
BE1	Blood	Complete	UK	-	NZ_LN877293.1
BF8	Infection site	Contig	Hungary	1985	NZ_LGTH00000000.1
BFG-1	Blood	Complete	USA	1985	NZ_CP081922.1
BFR_KZ01	Purulent sample	Contig	Kazakhstan	Dec-2018	JGDW00000000
BFR_KZ02	Wound	Contig	Kazakhstan	Nov-2018	PDCW01000000
BFR_KZ03	Faeces	Contig	Kazakhstan	Nov-2018	QSRG01000000
BFR_KZ05	Faeces	Scaffold	Kazakhstan	Mai-2019	NZ_JACENG00000000.1
BFR_KZ06	Faeces	Contig	Kazakhstan	Mai-2019	NZ_JACFSS00000000.1
BFR_KZ07	Faeces	Scaffold	Kazakhstan	Jul-2019	NZ_JACFST00000000.1
BFR_KZ08	Faeces	Scaffold	Kazakhstan	Aug-2019	NZ_JACFSU00000000.1
BFR_KZ09	Faeces	Contig	Kazakhstan	Oct-2019	NZ_JACFSV00000000.1
BFR_KZ10	Faeces	Scaffold	Kazakhstan	Oct-2019	NZ_JACFSW00000000.1
BFR_KZ11	Faeces	Contig	Kazakhstan	Nov-2019	NZ_JACFSX00000000.1
BFR_KZ12	Purulent sample	Contig	Kazakhstan	Nov-2019	NZ_JACFSY00000000.1
BOB25 ^a	Faeces	Complete	Russia	Mar-2013	NZ_CP011073.1
CCUG4856T	Faeces	Complete	UK	1955	NZ_CP036555.1
Plasmid pBF9343					NZ_CP036556.1
CFPLTA004_1B	Faeces	Contig	USA	2018	NZ_VOHY00000000.1
CL03T00C08	Faeces	Scaffold	USA	-	NZ_AGXK00000000.1
CL03T12C07	Faeces	Complete	USA	Mar-2009	CP072257.1
CL04T03C20	Faeces	Scaffold	USA	Mar-2009	NZ_PDCU00000000.1
CL05T00C42	Faeces	Contig	USA	-	NZ_AKBY00000000.1
CL05T12C13	Faeces	Scaffold	USA	-	NZ_AGXP00000000.1
CL07T00C01	Faeces	Scaffold	USA	-	NZ_AGXM00000000.

					1
CL07T12C05	Faeces	Scaffold	USA	-	NZ_AGXN00000000.1
CM13	Faeces	Scaffold	USA	2015	NZ_PDCT00000000.1
					NZ_JABAGK00000000.1
COR2-248-WT-1	Faeces	Contig	Germany	Mar-2018	00.1
DCMOUH0017B	Faeces	Complete	Denmark	Jun-2006	NZ_CP036539.1
Plasmid pBFO17_1					NZ_CP036540.1
Plasmid pBFO17_2					NZ_CP036541.1
				2011/10/10	
DCMOUH0018B	Blood	Complete	Denmark		NZ_CP036542.1
plasmid pBFO18_1					
DCMOUH0042B	Faeces	Complete	Denmark	2008	NZ_CP036550.1
Plasmid pBFO42_1					NZ_CP036551.1
Plasmid pBFO42_2					NZ_CP036552.1
DCMOUH0067B	Faeces	Complete	Denmark	2009	NZ_CP036553.1
Plasmid pBFO67_1					NZ_CP036543.1
Plasmid pBFO18_2					NZ_CP036544.1
plasmid pBFO18_3					NZ_CP036545.1
DCMOUH0085B	Faeces	Complete	Denmark	2010	NZ_CP037440.1
DCMSKEJBY0001B	Faeces	Complete	Denmark	Aug-2013	NZ_CP036546.1
Plasmid pBFS01_1					NZ_CP036547.1
Plasmid pBFS01_2					NZ_CP036548.1
Plasmid pBFS01_3					NZ_CP036549.1
DK1985	Blood	Complete	Denmark	1985	NZ_CP044428.1
DS-166 ^a	Tissue	Contig	USA	-	NZ_JGDD00000000.1
DS-208 ^a	Tissue	Contig	USA	-	NZ_JGDE00000000.1
Ds-233 ^a	Tissue	Contig	USA	-	NZ_JGDG00000000.1
DS-71 ^a	Tissue	Contig	USA	-	NZ_JGDF00000000.1
FDAARGOS_763	Faeces	Complete	USA	-	NZ_CP054003
Plasmid unnamed1					NZ_CP054001.1
					NZ_CP054002.1
Plasmid unnamed2					NZ_CP054004.1

Plasmid unnamed3					
GUT04	Faeces	Complete	Korea	Aug-2019	NZ_CP043610.1
Plasmid unnamed1					NZ_CP043609.1
HAP130N_1B	Faeces	Contig	USA	2018	NZ_WCIJ00000000.1
HAP130N_2B	Faeces	Contig	USA	2018	NZ_VOIS00000000.1
HAP130N_3B	Faeces	Contig	USA	2018	NZ_WCIE00000000.1
HCK-B3	Faeces	Contig	China	Apr-2017	NZ_QRBX00000000.1
HMW 610	Faeces	Scaffold	USA	-	NZ_AGXQ00000000.1
HMW 615	Faeces	Scaffold	USA	-	NZ_AGXR00000000.1
HMW 616	Faeces	Scaffold	USA	-	NZ_ALOB00000000.1
I1345 ^a	L colon sample	Contig	USA	Dec-2012	NZ_JGEW00000000.1
J1101437_171009_F3	Faeces	Scaffold	USA	Oct-2017	NZ_JADNMB01000001.1
J-143-4 ^a	Tissue	Contig	USA	-	NZ_JGDH00000000.1
J38-1 ^a	Tissue	Contig	USA	-	NZ_JGDV00000000.1
JIM10	Faeces	Contig	Russia	Mar-2013	NZ_MBRB00000000.1
KLE1257	Faeces	Scaffold	USA	-	NZ_LTZL00000000.1
KLE1758	Blood	Scaffold	USA	-	NZ_LTZK00000000.1
Korea 419 ^a	Tissue	Contig	USA	-	JMZX02000000
MC1	Faeces	Scaffold	USA	-	NZ_CAAKNW00000000.1
MGYG-HGUT-00236	Blood	Scaffold	UK	-	NZ_CABJEQ00000000.1
MGYG-HGUT-01337	Blood	Scaffold	UK	-	NZ_CABKOU00000000.1
NCTC9343	Diarrhea;Abscesses	Contig	UK	-	NZ_UFTH00000000.1
NCTC 9343	Appendix abscess	Complete	UK	2017	NC_003228.3
O:21	Blood	Scaffold	Sweden	1995	NZ_LMAC00000000.1
OF01-1	Faeces	Scaffold	China	Dec-2013	NZ_QSDG00000000.1
OF03-7	Faeces	Scaffold	China	Dec-2013	SFHX01000000
OF05-11AC	Faeces	Scaffold	China	Jan-2014	NZ_QSWE00000000.1
OF05-13AC	Faeces	Scaffold	China	Jan-2014	NZ_QSWB00000000.1

OM02-11	Faeces	Scaffold	China	Nov-2013	NZ_QSVT00000000.1
OM02-9	Faeces	Scaffold	China	Nov-2013	NZ_QSVD00000000.1
OM04-9BH	Faeces	Scaffold	China	Dec-2013	NZ_QSUS00000000.1
OM06-1	Faeces	Scaffold	China	Dec-2013	NZ_QSUD00000000.1
OM06-30AC	Faeces	Scaffold	China	Dec-2013	NZ_QSTV00000000.1
Q1F2	Faeces	Complete	Hong Kong	Feb-2013	NZ_CP018937.1
Plasmid Q1F2-p1					NZ_CP018938.1
Plasmid Q1F2-p2					NZ_CP018939.1
S13 L11 ^a	Tissue	Contig	USA	Nov-2011	NZ_JGDI00000000.1
S14 ^a	Infection	Complete	Hungary	2014	NZ_CP012706.1
S23 R14 ^a	Tissue	Contig	USA	Feb-2012	NZ_JGDX00000000.1
S23L17 ^a	Tissue	Contig	USA	Feb-2012	NZ_JHEF00000000.1
S23L24 ^a	Tissue	Contig	USA	Feb-2012	NZ_JGEL00000000.1
S24L15 ^a	Tissue	Contig	USA	Feb-2012	NZ_JGEM00000000.1
S24L26 ^a	Tissue	Contig	USA	Feb-2012	NZ_JGEN00000000.1
S24L34 ^a	Tissue	Contig	USA	Feb-2012	NZ_JGEO00000000.1
S36L11 ^a	Patient sample with septicemia	Contig	USA	Jul-2012	NZ_JGDI00000000.1
S36L12 ^a	Patient sample with septicemia	Contig	USA	Jul-2012	NZ_JGEP00000000.1
S36L5 ^a	Tissue	Contig	USA	Jul-2012	NZ_JGEQ00000000.1
S38L3 ^a	Healthy biopsy Tissue	Contig	USA	Aug-2012	NZ_JGEV00000000.1
S38L5 ^a	Appendix	Contig	USA	Aug-2012	NZ_JGER00000000.1
S6L3 ^a	Faeces	Contig	USA	Aug-2011	NZ_JGDY00000000.1
S6L5 ^a	Faeces day 7 after weanning	Contig	USA	Aug-2011	RCXN01000000
S6L8 ^a	Purulent sample	Contig	USA	Aug-2011	NZ_JGES00000000.1
S6R6 ^a	Purulent sample	Contig	USA	Aug-2011	NZ_JGDZ00000000.1
S6R8 ^a	Appendix abscess	Contig	USA	Aug-2011	NZ_JGVE00000000.1
TF05-31	Infection site	Scaffold	China	Aug-2013	NZ_QSSJ00000000.1
TF09-1	Faeces	Scaffold	China	Feb-2014	NZ_QSRG00000000.1

TL139C_1B	Faeces	Contig	USA	2018	NZ_VOHX00000000.1
TL139C_2B	Faeces	Contig	USA	2018	NZ_WCII00000000.1
TM08-15	Faeces	Scaffold	China	Feb-2014	NZ_QSOR00000000.1
US326	Faeces (Healthy Adult)	Scaffold	USA	2015	NZ_PDCS00000000.1
YCH46	Faeces (Healthy Adult)	Complete	Japan	-	NC_006347.1
YCH46 strain 82A12	Faeces (Healthy Adult)	Contig	Japan	-	NZ_UYXF00000000.1

^a Enterotoxigenic *Bacteroides fragilis* (ETBF) (Science, 2013).