Corso di dottorato di ricerca in:

"Alimenti e salute umana"

Ciclo XXXV

Titolo della tesi

"Genetic approach for in depth taxonomic characterization and safety traits of probiotic and ancient bacteria"

Dottorando
Andrea Colautti

Supervisore
Lucilla Iacumin

Anno 2023

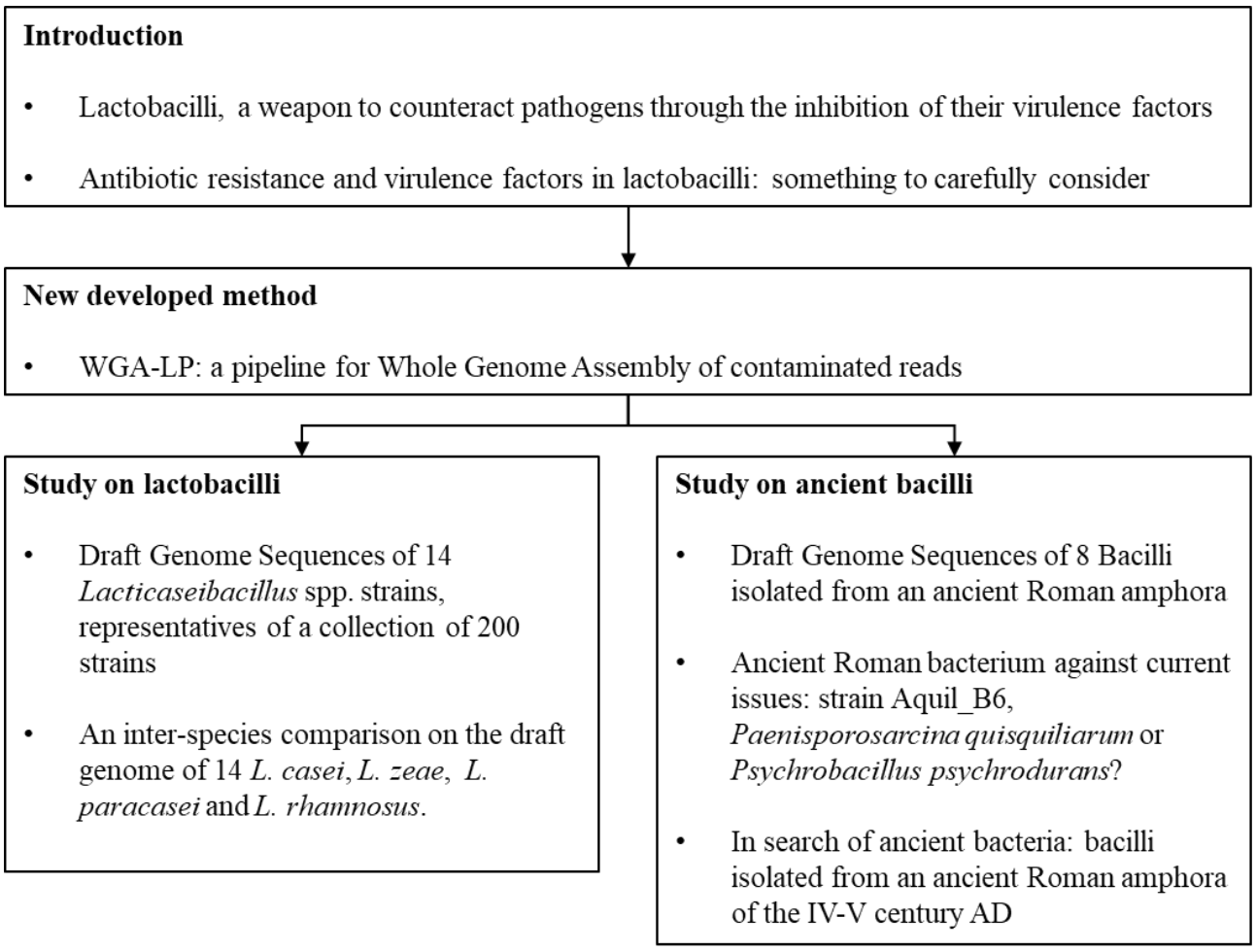
Introduction and thesis layout

This study has focussed on the genetic aspects of bacteria with potential probiotic applications, specifically, the former genera *Lactobacillus* spp. and *Bacillus* spp. In fact, due to their importance for use as probiotics or in functional foods, nowadays many researches on these microbial groups are conducted. However, in addition to analyze their technological characteristics, regulatory agencies have established that, before being used in products intended for human consumption, each microorganism have to be correctly classified at a taxonomic level and tested for the presence of potential human safety risk factors. In this regards, genetic characterization via whole genome sequencing appears to be one of the most effective classification techniques, as well as allowing the identification of potentially phenotypically unobservable genetic factors of virulence and resistance. The first introductory part of the thesis therefore focused on the scientific evidence of the beneficial effects, as well as potential risk factors, of the former genus *Lactobacillus* spp. In particular, in the first chapter entitled "Lactobacilli, a weapon to counteract pathogens through the inhibition of their virulence factors", the collected scientific literature demonstrated the ability of different species of this genus to reduce and counteract the effects of virulence factors of various pathogens, such as *Aggregatibacter actinomycetemcomitans*, *Bacillus cereus*, *Campylobacter jejuni*, *Candida albicans*, *Chlamydia trachomatis*, *Clostridium* spp., *Enterococcus faecalis*, *Escherichia coli*, *Gardnerella vaginalis*, *Helicobacter* spp., *Klebsiella* spp., *Listeria monocytogenes*, *Neisseria gonorrhoeae*, *Pseudomonas* spp., *Prevotella bivia*, *Salmonella* spp., *Serratia marcescens*, *Staphylococcus aureus*, *Streptococcus* spp., *Trichomonas vaginalis*, and *Yersinia* spp.. The collected evidence are very important as they lay the foundations for possible further studies that demonstrate the efficacy of Lactobacilli against pathogens even for direct application in humans. In fact, selected strains could fight infections alongside with current antibiotic treatments, increasing their effectiveness. One of the main problems of modern medicine is the phenomenon of the spread of antibiotic resistance, and without the discovery of new antibiotic substances, it is imperative to make current therapies increasingly effective, and to reduce the use of antibiotics as much as possible. In this regard, the regular intake of Lactobacilli, in addition to the already proven probiotic properties and proven anti-virulence effects, could be a valid aid in preventing the onset of infections. However, as previously reported, in addition to the innumerable beneficial effects, it is of fundamental importance to identify the presence of risk factors in these bacteria. This issue is dealt in the second chapter "Antibiotic resistance and virulence factors in lactobacilli: something to carefully consider". In this literature review, all possible resistance and virulence factors reported for lactobacilli have been summarized. Although they are generally considered safe, indeed, the large number of risk factors identified in the numerous works in the literature suggest caution in their use especially in frail subjects with pre-

existing or cronic diseases and the importance of a preliminary screening before their use. However, to date, there are no guidelines to follow these assessments, and there are no specific databases with the list of possible risk factors to search for. One of the main and most effective techniques currently available for analyzing and identifying the presence of risk or beneficial factors is the *in silico* study of the genome of microorganisms. With the reduced costs of Whole Genome Sequencing (WGS) technology, it is in fact possible to obtain the entire genetic sequence of the microorganism under analysis and to study their genetic characteristics quickly and effectively. Several pipelines are currently available for the assembly of the reads obtained from the sequencing process, which however do not take care of the pre- and post-assembly quality. In the chapter "WGA-LP: a pipeline for Whole Genome Assembly of contaminated reads" a new pipeline was therefore developed with the aim of combining the assembly phase with the procedures for verifying the quality of both the raw reads and the assembly obtained, paying particular attention to the presence of possible contaminants. In fact, many genomes found in databases are contaminated by foreign DNA. It is therefore important to carefully verify the genetic sequences before their divulgation to avoid the propagation of errors in the scientific community. Thanks to this pipeline, the genomes of 14 lactobacilli strains representative of a collection of 200 strains in the chapter “Draft Genome Sequences of 14 *Lacticaseibacillus* spp. strains, representatives of a collection of 200 strains”, which were later characterized in the chapter “An inter-species comparison on the draft genome of 14 *L. casei*, *L. zeae, L. paracasei* and *L. rhamnosus”* were assembled. It was therefore possible to carry out a more precise taxonomic classification than that carried out in the previous works, managing to obtain precise identifications even for some strains that were not identified with certainty, considering the close related species under study. It was also possible to verify the presence of potential risk factors, in this case identified on a plasmid of a strain. At the same time, the genetic characterization was also performed on several strains of bacilli isolated from an ancient Roman amphora. From the dehydrated contents of the amphora dating back to the 4th-5th century AD, found still intact and sealed during excavations in the city of Aquileia (UD), it was in fact possible to isolate still viable spores of bacteria of the genus *Bacillus*. After an initial genetic selection performed with classical molecular methods, 8 strains with a unique genetic fingerprint profile were selected. After WGS and a first taxonomic identification, the sequences obtained were published in the work "Draft Genome Sequences of Eight Bacilli Isolated from an Ancient Roman Amphora". All the strains resulted phylogenetically very close to each other, and once again, the possibility of comparing the whole genomes allowed to obtain a precise identification of the species, not achievable with methods such as 16S rRNA gene sequencing. However, during the taxonomic identification of a strain, inconsistencies emerged in the genomes of the reference strains with which it clustered. In the chapter

“Ancient Roman bacterium against current issues: strain Aquil_B6, *Paenisporosarcina quisquiliarum* or *Psychrobacillus psychrodurans*?” emerged an error in the genetic sequence of the reference strain *Paenisporosarcina quisquiliarum* SK55, which was then resequenced together with *Psychrobacillus psychrodurans* strains DSM 11713 and DSM 30747. Thanks to the new sequences of higher quality than those previously available, it was possible to shed light on the species to which the Aquil_B6 strain belongs, coming to a succesful and unique identification. Finally, in the chapter "In search of ancient bacteria: bacilli isolated from an ancient Roman amphora of the IV-V century AD", the remaining strains isolated from the amphora are analysed. From an in-depth taxonomic analysis, it emerged that strains Aquil_B1 and Aquil_B8 for ANI and dDDH values were a potential new species. Also, in the functional annotation, several possible antibiotic resistance factors emerged, testifying to the diffusion of these genetic traits in bacteria that had remained isolated since times in which antibiotics were not widely used. A brief graphic summary of the thesis is shown below.

Thesis Layout



Index

Chapter 1: Lactobacilli, a weapon to counteract pathogens through the inhibition of their virulence factors

Colautti Andrea, Orecchia Elisabetta, Comi Giuseppe, Iacumin Lucilla*

Department of Agricultural, Food, Environmental and Animal Science, University of Udine, via Sondrio 2/A, 33100 Udine, Italy

*Corresponding author:
Lucilla Iacumin,
Department of Food Science,
via Sondrio 2/A, 33100 Udine, Italy,
tel: +390432558126, Fax: +390432558130,
E-mail: lucilla.iacumin@uniud.it

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology®

MINIREVIEW

Lactobacilli, a Weapon to Counteract Pathogens through the Inhibition of Their Virulence Factors

Andrea Colautti,[a] Elisabetta Orecchia,[a] Giuseppe Comi,[a] Lucilla Iacumin[a]

[a]Department of Agriculture, Food, Environmental and Animal Sciences, University of Udine, Udine, Italy

ABSTRACT To date, several studies have reported an alarming increase in pathogen resistance to current antibiotic therapies and treatments. Therefore, the search for effective alternatives to counter their spread and the onset of infections is becoming increasingly important. In this regard, microorganisms of the former *Lactobacillus* genus have demonstrated the ability to reduce the virulence of pathogens. In addition to the production of bioactive substances, self- and coaggregation, and substrate competition, lactobacilli influence gene expression by downregulating genes associated with the virulence of pathogens. As demonstrated in many *in vivo* and *in vitro* trials, lactobacilli counteract and inhibit various virulence factors that favor pathogens, including the production of toxins, biofilm formation, host cell adhesion and invasion, and downregulation of virulence genes linked to quorum sensing. The aim of this review is to summarize current studies on the inhibition of pathogen virulence by lactobacilli, an important microbial group well known in the industrial and medical fields for their technological and probiotic properties that benefit human hosts with the potential to provide an important aid in the fight against pathogens besides use of the current therapies. Further research could lead to the identification of new strains that, in addition to alleviating adverse effects, could improve the efficacy of antibiotic therapies or play an important preventive role by reducing the onset of pathogen infections if regularly taken.

KEYWORDS lactobacilli, virulence, probiotics, pathogen suppression

Lactobacilli, the term used in this work to refer to the former *Lactobacillus* genus (1), are lactic acid bacteria with fundamental roles in modern society and economies and are essential in the production and conservation of many food and feed products. Owing to their long history of safe use and their fermentative and bioprotective abilities, which ensure the quality and safety of products, they have received the designations of generally recognized as safe by the Food and Drug Administration and qualified presumption of safety by the European Food Safety Authority (EFSA) (2, 3). Due to their properties, several strains of this group have been identified as probiotics, defined by FAO and WHO as "live microorganisms which when administered in adequate amounts confer a health benefit on the host" (4, 5), and their inactivated cells or their cell-free supernatants (CFS) hosting numerous beneficial components are also considered postbiotics, defined as "preparation of inanimate microorganisms and/or their components that confers a health benefit on the host" (6). They are also part of the human natural bacterial flora, in which they have a regulatory role in protecting hosts against colonization bypathogens and exert beneficial effects, such as increasing and improving nutrient assimilation during digestion or stimulating host tissues (7). Prolonged consumption of these bacteria leads to modification of the human gastrointestinal microbial flora, thus stimulating the immune system and decreasing pathogen adhesion (8). Owing to the interconnection between the gastrointestinal tract and the central nervous system, known as the gut-brain axis, these effects also arise from the production of signaling molecules with brain modulation abilities (9, 10).

Editor Laurie E. Comstock, University of Chicago

Address correspondence to Lucilla Iacumin, lucilla.iacumin@uniud.it.

The authors declare no conflict of interest.

Published 26 October 2022

1.1 Abstract

To date, several studies have reported an alarming increase in pathogen resistance to current antibiotic therapies and treatments. Therefore, the search for effective alternatives to counter their spread and the onset of infections is becoming increasingly important. In this regard, microorganisms of the former *Lactobacillus* genus have demonstrated the ability to reduce the virulence of pathogens. In addition to the production of bioactive substances, self- and coaggregation, and substrate competition, lactobacilli influence gene expression by downregulating genes associated with the virulence of pathogens. As demonstrated in many *in vivo* and *in vitro* trials, lactobacilli counteract and inhibit various virulence factors that favour pathogens, including the production of toxins, biofilm formation, host cell adhesion and invasion, and downregulation of virulence genes linked to quorum sensing. The aim of this review is to summarize current studies on the inhibition of pathogen virulence by lactobacilli, an important microbial group well known in the industrial and medical fields for their technological and probiotic properties that benefit human hosts with the potential to provide an important aid in the fight against pathogens besides use of the current therapies. Further research could lead to the identification of new strains that, in addition to alleviating adverse effects, could improve the efficacy of antibiotic therapies or play an important preventive role by reducing the onset of pathogen infections if regularly taken.

Keywords

lactobacilli, virulence, probiotics, pathogen suppression

1.2 Introduction

Lactobacilli, the term used in this work to refer to the former *Lactobacillus* genus (1), are lactic acid bacteria with fundamental roles in modern society and economies and are essential in the production and conservation of many food and feed products. Owing to their long history of safe use and their fermentative and bioprotective abilities, which ensure the quality and safety of products, they have received the designations of generally recognized as safe by the Food and Drug Administration and qualified presumption of safety by the European Food Safety Authority (EFSA) (2, 3). Due to their properties, several strains of this group have been identified as probiotics, defined by FAO and WHO as "live microorganisms which when administered in adequate amounts confer a health benefit on the host" (4, 5), and their inactivated cells or their cell-free supernatants (CFS) hosting numerous beneficial components are also considered postbiotics, defined as "preparation of inanimate microorganisms and/or their components that confers a health benefit on the host" (6). They are also part of the human natural bacterial flora, in which they have a regulatory role in protecting hosts against colonization by pathogens and exert beneficial effects, such as increasing and improving nutrient assimilation during digestion or stimulating host tissues (7). Prolonged consumption of these bacteria leads to modification of the human gastrointestinal microbial flora, thus stimulating the immune system and decreasing pathogen adhesion (8). Owing to the interconnection between the gastrointestinal tract and the central nervous system, known as the gut-brain axis, these effects also arise from the production of signalling molecules with brain modulation abilities (9, 10). Lactobacilli are also effective in the prevention and treatment of gastrointestinal and urogenital tract diseases because of their antimicrobial properties (11, 12) and confer numerous beneficial effects, such as alleviating lactose intolerance, reducing blood cholesterol and incidence and progression of cancer, stimulating immunity, and preventing and treating diarrheal diseases, stomach ulcers, and infectious diseases (13, 14). Furthermore, lactobacilli inhibit pathogen growth through nutrient subtraction, competition for substrate, and the production of molecules such as bacteriocins, enzymes, organic acids, and hydrogen peroxide (15). Other important mechanisms include the ability to self-aggregate and coaggregate, which allow lactobacilli to adhere to each other or other microbial species. These adhesive properties provide lactobacilli with the ability to adhere to the mucosa, thereby limiting pathogen adhesion and creating a microenvironment in which their strict proximity allows the increase of inhibitory effects of the secreted substances (16). In addition to these well-known properties, lactobacilli inhibit various virulence genes encoding transacting proteins associated with infective mechanisms, which are fundamental in bacterial virulence, as reviewed in Table 1. Among these mechanisms, one of the most important is the quorum sensing (QS) system, which leads to the production of different chemical molecules, named autoinducers, which alter gene expression.

Through these signal-response systems, different bacteria coordinate their behaviours on a population scale, acting as multicellular organisms (17). QS systems regulate many microbial pathways, including biofilm formation, sporulation, antibiotic synthesis, induction of virulence factors, host infection, and bacteriocin synthesis. Autoinducer 2 (AI-2), produced by the LuxS enzyme (*luxS* gene), is of particular interest because it is associated with the expression of genes involved in pathogen motility, adhesion, and internalization. AI-2 also plays a fundamental role in biofilm formation, a common feature among pathogenic species that increases their adhesion to surfaces, provides them with nutrients, and confers resistance to external factors, thus making bacteria more virulent and resistant to antibiotic treatments (18–20). Moreover, antiviral activity, a property of particular interest in medical applications, has been observed in specific strains of lactobacilli and might be used to prevent viral adhesion and propagation (21). Pathogenic bacteria are an important threat to human health, as they represent 4 of the top 10 causes of death worldwide (22). Currently, infections are treated mainly with antibiotics, whose discovery dates to the first half of the 20th century. However, the extensive and prolonged use of these substances has led to a natural evolutionary phenomenon of adaptation that has contributed to the spread of antibiotic resistance (23). Consequently, infections have become more difficult because antibiotics have become less effective in counteracting pathogens, thus enabling their survival and even replication in the presence of therapeutic levels of drugs. If no action is taken, multidrug-resistant pathogens have been expected to cause 10 million deaths by the year 2050. Therefore, identifying new effective methods will be critical to counteract the spread of pathogens and simultaneously decrease the use of antibiotics (24) in medical and zootechnical fields (25). The present review summarizes available data from original studies reporting the effectiveness of lactobacilli in counteracting the virulence of pathogenic species such as *Aggregatibacter actinomycetemcomitans*, *Bacillus cereus*, *Campylobacter jejuni* (Cj), *Candida albicans*, *Chlamydia trachomatis*, *Clostridium* spp., *Enterococcus faecalis*, *Escherichia coli* (*Ec*), *Gardnerella vaginalis*, *Helicobacter* spp., *Klebsiella* spp., *Listeria monocytogenes* (Lm), *Neisseria gonorrhoeae*, *Pseudomonas* spp., *Prevotella bivia*, *Salmonella* spp., *Serratia marcescens*, *Staphylococcus aureus* (*Sa*), *Streptococcus* spp., *Trichomonas vaginalis*, and *Yersinia enterocolitica*, as summarized in Fig. 1.

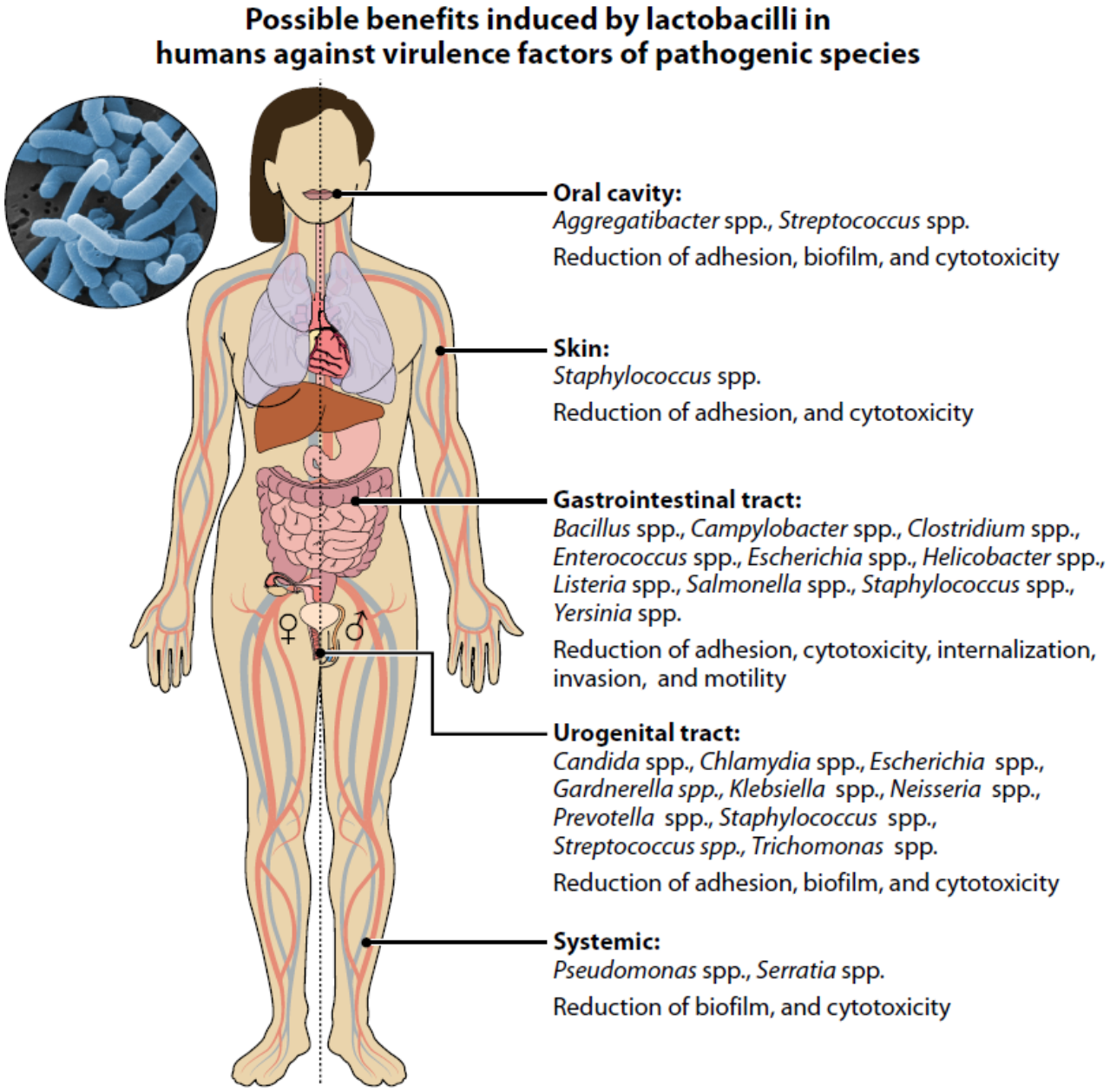


Fig 1 Possible benefits induced by lactobacilli in humans against virulence factors of pathogenic species.

Table 1: Summary of virulence genes affected by Lactobacilli

Gene	Protein	Function	Reference
Listeria monocytogenes			
fbp	Fibronectin binding protein	Adhesion to epithelial cells	114
flaA	Flagellin	Motility	28
hly	Hemolysin listeriolysin O (LLO)	Survival inside macrophages	27
iap	Invasion-associated protein	Invasion of epithelial cells	114
plcA plcB	PlcA PlcB phospholipases	Survival inside macrophages	27
prfA		Transcriptional activator of *hly* and *plc* genes	29
sigB		Stress response regulon	31
	Autolysin amidase (Ami)	Bacterial adhesion on enterocytes	29
	Actin-polymerizing protein (ActA)	Required for actin polymerization allowing intracytoplasmatic movement	30
	Internalin A (InlA) Internalin B (IlnB)	Adhesion and internalization inside enterocytes	27
	Listeria adhesion protein (LAP)	Bacterial adhesion on enterocytes	29
Salmonella spp.			
avrA	AvrA	Inhibition of innate immunity	56
hilA	HilA	Regulation of *Salmonella* Pathogenicity Island 1 gene expression	48
hilC hilD	HilC HilD	Transcriptional regulators of *hilA*	48
invH	Outer membrane lipoprotein InvH	Facilitates the rtanslocation of proteins including SipC from the cytoplasm to the membrane	53
nmp	Outer membrane-associated protein	Bacterial porin formation	114
prgK	*PrgK* periplasmatic protein	Type III secretion system	49
sip	Sip effector protein	Induction of inflammation response	52
sop	*Salmonella* outer Protein B	Lipid phosphatase critical in enteropathogenicity	50
sptP	SptP effector protein	Recovery of the host cytoskeleton after the infection	55
spv		Promoter of the virulence genes of non-typhoid *Salmonella* serovars	51
ssrB	SsrB	Activation of genes needed for intracellular survival	57
Campylobacter jejuni			
cadF	Outer membrane protein CadF	Adhesion to intestinal peithelial cells	76

cdt	Cytolethal distending toxin	Toxin composed by three subunits, involved in cell adhesion and inhibition of cell division	76
cia	*Campylobacter* invasion antigen B	Invasion potential	76
fla	Flagellin	Motility and colonization	76
flh	Flagellin	Motility and colonization	76
luxS	LuxS enzyme	Production of Autoinducer 2 AI-2	79
Escherichia coli			
eaeA	Intimin	Attachment to cell surface	86
fliC	Flagellin	Motility	96
hly	Enterohemolysin and α-hemolysin	Toxins with hemolytic activity	87
ler	LEE1- encoded regulator	Transcriptional activator of LEE genes	94
luxS	LuxS enzyme	Production of Autoinducer 2 AI-2	97 98
qseA	QseA effector protein	*LEE1* gene activator	95
stx	Shiga-like toxin Stx	Toxing causing diarrhoea and other disorders	89
tir	Translocated Intimin Protein	Adhesion to epithelial cells	93
	Adhesins	Adhesion on both abiotic and celle surface	91
	Intimin receptor EspE	Type III secretion system that allows A/E lesions	92
***Clostridium* spp.**			
luxS	LuxS enzyme	Production of Autoinducer 2 AI-2	125
tcdA	Enterotoxin A	Toxin wich causes diarrhoea and intestinal damage	119 120
tcdB	Toxin B	Toxin with strong cytotoxic effect	119, 120
txeR	σ Factor	Induces RNA polymerase to recodgnize the promoters of *tdc* genes	121
Staphylococcus aureus			
agr		QS system which regulates virulence factors	130
ica		Biofilm formation	137
mecA		Methicillin reisitance	136
sae		Regulatory locus wich activates the production of different exoproteins	131
sbi	Immunoglobuli-binging protein	Binding to immunoglobulin G and blood coagulation	135
sea	Enterotoxin A	Food poisoning	132
spa	Protein A	Inhibition of phagocytosis	135
ssl1	*Stpahylococcus* superantigen-like protein (SSL 1)	Inhibition of metalloproteases	134
tst	Toxic shock syndrome toxin-1 (TSST-1)	Superantigen wich causes organ dysfunctions associated with high mortality rate	133

***Helicobacter* spp.**			
cagA	CagA cytotoxin	Alteration of intracellular signal transduction	148
fla	Flagellin	Motility	149
vacA	VacA cytotoxin	Fusion between endosomes and lysosomes in eukatyotic cells	148
Pseudomonas spp.			
exo	Cytotoxins belonging to the type III effector proteins family	Toxins wich cause different damage to the host	157
fleSR	Flagellin	Flagella necessary to swimming/swarming motility	158
lasI/R	LasI/R protein	QS system which regulates virulence factors	162
ndvB		Biofilm formation	157
pil	Pilin	Type IV pili necessary to twitching motility	158
rhI/R	RhI/R protein	QS system which regulates virulence factors	162
Klebsiella pneumoniae			
sugE		Biofilm formation	163
treC			163
Streptococcus spp.			
ftf	Fructosyltransferase	Adhesion	168
gtf	GTF glucotransferase	Production of exopolysaccharides	167
luxS	LuxS enzyme	Production of Autoinducer 2 AI-2	171
sag	Streptolysin S	Toxin wich causes erythrocytes lysis	177
tft	GTF glucotransferase	Production of exopolysaccharides	167
Neisseria gonorrhoeae			
	Major outer protein porin PorB	Suppression of neutrophil oxidative burst and apoptosis	187
	N. gonorrhoeae lipooligosaccharide LOS	Adhesion and invasion of the host cells	187
	Opacity proteins (Opa)	Colonization of the mucosal eputhelium	187
	Pilin	Type IV pili for twitching motility, immune evasion and colonisation	187
Trichomonas vaginalis			
	Lipophosphoglycan	Adherence factor	186
Gardnerella vaginalis			
sld	Sialidase	Adhesion to cells and surfaces	188
vly	Vaginolysin	Inhibition of immune response	
Candida albicans			
ALS3	Adhesins	Adhesion properties	195
BCR1		Biofilm formation	195
CPH1		Biofilm formation	195
ECE1		Yeast to hyphal morpogenesis	196
EFG1		Biofilm formation	195
HWP1	Adhesins	Adhesion properties	195

Msp1	Major peptidoglycan hydrolase	Chitin hydrolysis	204
Saps		Hydrolytic enzymes	196
TEC1		Biofilm formation	195
	CDR1, CDR2, MDR1 proteins	Resistance to drugs and immune system	195
Aggregatibacter actinomycetemcomitans			
LtxA	Leukotoxins	Induces the death of leukocytes	220
CdtB	Cytholetal distending toxin	diarrheal disease-causing toxin	

1.3.1 *Listeria monocytogenes*

Listeria monocytogenes (*Lm*) is the etiological agent of listeriosis, a severe foodborne disease with a low incidence rate but a high mortality rate that poses a serious public health concern (26). Internalization of this pathogen occurs via invasion of macrophages and nonphagocytic cells, a capability conferred by the internalin proteins *InlA* and *InlB*, while the production of hemolysin listeriolysin O (LLO) and PlcA and PlcB phospholipases, encoded by the *hly* and *plc* genes, respectively, enables macrophage survival (27). The presence of *Listeria* adhesion protein (LAP) and autolysin amidase Ami, which enhance bacterial adhesion, *prfA* transcriptional activator, ActA actin polymerization protein, *sigB* stress response factor, and flagellin, encoded by *flaA* gene, all contribute to *Lm* virulence (28–31). Several studies have reported the reduction of all of these virulence factors (Table S1 in the supplemental material). *In vitro* trials have revealed that lactobacilli, through the production of organic acids and proteinaceous molecules and their interaction with mucosal epithelial cells, significantly decreased inflammation during the invasion of *Lm* (32). Coculture with *Lactiplantibacillus plantarum* significantly decreased *Lm* virulence toward HT-29 cells (33). On Caco-2 cells, *Lpb. plantarum* and *Lacticaseibacillus rhamnosus* coinoculation significantly reduced the *Lm* survival ratio under simulated digestion, thus inhibiting cell adhesion and invasion and downregulating the *sigB*, *hly*, *inlA*, *inlB*, and *prfA* genes (34, 35). This property was also observed for *Limosilactobacillus reuteri*, *Limosilactobacillus fermentum*, and *Lpb. plantarum* with lower LLO production, epithelial E-cadherin-binding ability, and expression of virulence genes, while in an *in vivo* trial, these strains increased survival of *Galleria mellonella* inoculated with lethal doses of *Lm* (36). In addition, preexposure to bioengineered *Lacticaseibacillus casei* and Lacticaseibacillus paracasei preserved tight barrier junction integrity and decreased *Lm*-mediated cytotoxicity and adhesion, whereas these effects were not observed on *Lm* already attached to Caco-2 cells (37, 38). Other *in vivo* studies confirmed the antilisterial activities of lactobacilli. In murine models, the administration of *Lcb. paracasei* and *Lcb. casei* systematically decreased the dissemination of *Lm* (39), whereas *Latilactobacillus sakei* 2a lowered lesions and edema of the intestinal villi (40).

Levilactobacillus brevis reduced the propagation and dispersion of *Lm* in the intestines, spleen, and liver without affecting neutrophils and lymphocyte values (41). In infected chickens, supplementation with *Lactobacillus acidophilus* and *Lpb. plantarum* attenuated *Lm* adhesion, pore formation, and invasion, downregulating the expression of LLO, *InlA*, *InlB*, Ami, and flagellin. Moreover, a decreased load of *Lm* in the cecum, skin, liver, and spleen, a decrease in serum cytokines, and an upregulation of antiinflammatory- related genes were observed (42). In addition, *Lm* cocultured with bacteriocin- producing *Llb. sakei* 1 resulted in diminished hemolytic activity (43, 44), thus indicating the effectiveness of lactobacilli in preventing *Lm* adhesion to abiotic surfaces (45, 46).

1.3.2 *Salmonella* spp.

Salmonella enterica (*Slm*) is a pathogen that affects both humans and animals. Septicemia and enteric fever are common clinical manifestations of serovars Typhi and Paratyphi, whereas bacteremia is typical of nontyphoidal *Salmonellae*, such as *S. enterica* serovar Typhimurium (*Slm*T), Enteritidis (*Slm*E), Heidelberg (*Slm*H), and Javiana (*Slm*J) (47). *Salmonella* pathogenicity islands (SPI) group *hilA*, *hilC*, and *hilD* invasion genes (48) and *prgK*, which are associated with type III secretion system 1 (T3SS1) and T3SS2 systems (49), as well as sop genes, which are important in enteropathogenesis (50). The virulence traits of nontyphoid *Salmonella* serovars are also enhanced by the spv plasmidic gene (51). The *invH* gene promotes tissue invasion both *in vivo* and *in vitro* and is related to the expression of the *sip* gene, which is involved in host translocation (52, 53). During infection, *Slm* invades macrophages and dendritic and epithelial cells (54), thus promoting survival and replication thanks to *avrA*, *sptP*, and *ssrB* genes (48, 55–57). Several studies have demonstrated that lactobacilli and their metabolites downregulate genes associated with *Slm* virulence (Table S2 in the supplemental material). *Lactobacillus bulgaricus*, *Lcb. paracasei*, and *Lcb. rhamnosus*, for example, downregulate the *sipA*, *sipB*, *sopB*, *spvB*, *hilA*, *hilD*, and *invH* genes in *Slm*E, *Slm*T, and *Slm*H (50), whereas *hilA* and *hilD* along with *hilC* and *sipC* are also downregulated by other probiotic lactobacilli (58). In *Slm*T-infected chickens administered lactobacilli, almost all SPI virulence genes (*hilA*, *hilC*, *hilD*, *sopB*, *sopD*, *sopE2*, *sipA*, *avrA*, and *sptP*, but not *sipC*) were downregulated, thus decreasing infection in the liver and spleen (59, 60). In addition, *Lbc. acidophilus* and *Lpb. plantarum* reduced the expression of the *invA*, *avrA*, *hilA*, *ssrB*, and *sopD* genes and the invasiveness of *Slm*T, thus altering the function of the type III secretion system (61, 62). A *Lbc. acidophilus* strain was also able to delay the internalization of *Slm*T, also altering its swimming motility (63). Other lactobacilli and their metabolites showed substantial antivirulence properties toward *Slm* in *in vivo* studies; for example, different *Lpb. plantarum* strains interfered with the growth and virulence of *Slm*T on Vero cells. These lactobacilli, which had higher ciprofloxacin resistance than the pathogen, significantly reduced its

adherence, invasion, and cytotoxicity (64). Preexposure of HT29 cells to live *Lbc. acidophilus*, *Lcb. rhamnosus*, and *Lcb. casei* decreased the induced cytotoxicity and the expression of virulence genes, particularly those related to the invasiveness of *Slm*J (65). Also, on thermally stressed Caco-2 cells, *Lcb. rhamnosus* reduced the severity of *Slm* infection (66). The adhesion of *Slm*T to the same cell line was inhibited by molecules secreted by lactobacilli, in particular lactic acid produced from *Lcb. casei* Shirota, *Lbc. acidophilus*, *Lcb. rhamnosus*, and *Lactobacillus. amylovorus*, whereas *Lactobacillus johnsonii* and *Lpb. plantarum* produced unknown inhibitory substances with anti-*Salmonella* activity (67). A bioengineered *Lcb. casei* strain overproducing conjugated linoleic acids (CLA) competitively excluded *Slm*T in a mixed culture and altered biofilm formation, adherence, and invasive activity toward INT-407 host cells, thus downregulating expression of the *invG*, *invH*, *prgK*, *hilA*, *hilC*, *hilD*, and *invF* genes (68, 69). Live lactobacilli cells and their CFSs show antivirulence effects against *Slm*. *Lcb. paracasei* CFS lowered *Slm*E adhesion to Caco-2 cells (70), whereas the CFS produced by *Lbc. acidophilus* induced the release of lipopolysaccharide in *Slm*T, a decrease in intracellular ATP correlated with bacterial death, bacterial membrane permeabilization, and increased sensitivity to sodium dodecyl sulfate (71). In a trial evaluating the expression of the *Slm*E *hilA-lacZY* transcriptional fusion, 24 h of incubation with spent medium from a *Lactobacillus* species strain isolated from poultry resulted in an absence of β-galactosidase activity. In comparison, *Slm*E, grown in *Slm*-spent medium, showed a 4-fold higher expression of *hilA* (72). Other properties of lactobacilli have been demonstrated *in vivo*. *Lcb. casei* inhibited the invasion and decreased the survival of *Slm*T in Caco-2 cells and mice, thus lowering the cecal colonization levels and the bacterial translocation rate to the spleen, liver, and mesenteric lymph nodes. In addition, administration of *Lcb. casei* to infected mice significantly delayed the occurrence of 100% animal mortality from 9 to 15 days (73). Pre-treatment with washed cells and CFS of *Ligilactobacillus salivarius*, *Lactobacillus delbrueckii* subsp. *delbrueckii*, and *Lpb. plantarum* inhibited *Slm*T attachment to the cecal mucus of infected chickens (74). The immune system modulation ability of lactobacilli was observed in *Slm*-infected mice, in which *Lacticaseibacillus zeae*, *Lpb. plantarum*, and *Lmb. reuteri* increased the proinflammatory cytokine response. This induced response was more effective with a combination of lactobacilli isolates than with a single strain (75).

1.3.3 *Campylobacter jejuni*

Campylobacter jejuni (*Cj*) is a commensal microorganism that is found in both domestic and wild animals and is responsible for campylobacteriosis, a severe foodborne diarrheal disease. Its virulence and survival in humans are linked to a variety of factors, including flagellum motility conferred by *fla* and *flh* genes, adhesion capacity conferred by *cia* and *cadF* genes, and cytolethal distending toxin

encoded by *cdtA*, *cdtB*, and *cdtC* genes, interfering with cell division (76). Lactobacilli, already recognized for their ability to relieve gastrointestinal symptoms caused by pathogenic infections, have been found to decrease *Cj* invasiveness (Table S3 in the supplemental material) (77). *In vitro* experiments revealed that the prolonged colonization of E12 cells with different lactobacilli attenuated *Cj* association, internalization, and translocation to the basolateral medium in transwells (78). On Caco-2 cells, various lactobacilli exhibited antagonistic effects against this pathogen, lowering the expression of genes involved in invasion (*ciaB*), motility (*flaA*, *flaB*, and *flhA*), and AI-2 production (*luxS*). These strains increased *Cj* macrophage phagocytosis and the expression of interferon-γ (IFN- γ), interleukin- 1β (IL-1β), IL-12p40, IL-10, and chemokines in macrophages (79). Similarly, the CFS of a genetically engineered *Lcb. casei* overexpressing the *mcrA* gene decreased *Cj* adhesion to, and invasion of, HD-11 and HeLa cells and altered the expression of *cadF*, *cdtB*, *ciaB*, and *flaB* genes (80). The expression of *ciaB* and *flaA* virulence genes in *Cj* was downregulated by *Lbc. acidophilus* CFS, according to real-time PCR (RT-PCR) analysis. The effect of the same strain has been tested on *luxS*-mutant *Cj* and downregulated only the *ciaB* gene, thereby suggesting an active role of *luxS* in the modulation of *Cj* virulence even when lactobacilli strains were added (81).

1.3.4 *Escherichia coli*

Although *Escherichia coli* (*Ec*) is commonly part of the commensal intestinal microbiota in both human and animal intestines, some opportunistic strains transmitted via the fecal-oral route can cause disease in humans. Pathogenic *Ec* can be classified as extraintestinal or diarrhoeagenic and can be further subdivided into different pathovars: enteropathogenic (EPEC), enterohemorrhagic (EHEC), enterotoxigenic (ETEC), enteroinvasive (EIEC), enteroaggregative (EAEC), Shiga toxin-producing (STEC), adherent invasive (AIEC), and diffusively adherent (DAEC) (82, 83). Whereas EIEC is an intracellular pathogen that invades and replicates within epithelial cells and macrophages, other pathogenic *Ec* strains interact with the epithelium through the expression of specific genes such as the *eaeA* gene, which regulates attachment to intestinal cells (84–86). An important virulence factor is the production of toxins, such as cell-associated enterohemolysin and α-hemolysin, encoded by *hlyA*, *hlyB*, *hlyC*, and *hlyD* genes in STEC (87). ETEC and EHEC are the main causes of enteric diseases in humans each year (88) owing to the ability of EHEC to produce verotoxin and Shiga-like toxins (Stx1 and Stx2) (89) and the ability of ETEC to produce toxins and adhesins (90, 91). EHEC has a pathogenicity island called locus of enterocyte effacement (LEE), which encodes gene regulators, adhesin, the type III secretion system, and proteins, including the translocated intimin receptor (*tir*) and Esp proteins that enhance adhesion to epithelial cells (92, 93). *LEE1*-encoded regulator (*ler*) activity is controlled by QS autoinducer 3 (AI-3) and by epinephrine and

norepinephrine hormones (94), whereas the *qseA* gene encodes the QseA effector protein, which directly activates the *LEE1* gene (95). EHEC is further characterized by the presence of a flagellum encoded by the *fliC* gene (96). Different lactobacilli and their metabolites alter the gene expression and consequently the virulence of *Ec* (Table S4 in the supplemental material). For example, *Lmb. reuteri* downregulated the epinephrine-mediated induction of *ler* in EHEC (94). CFS from *Lbc. acidophilus* supplementation in yogurt reduced the severity of infection and the attachment and colonization of EHEC and downregulated tumor necrosis factor-α (TNF-α) in infected mice. These effects were supported by RT-PCR, which detected a decrease in the expression of the $stxB_2$, *qseA*, *luxS*, *tir*, *ler*, *eaeA*, and *hlyB* genes (97). Another study found that CFS of the same strain reduced extracellular AI-2 concentrations and downregulated other virulence-associated genes (*tir*, *espA*, *fliC*, *espD*, *luxS*, *eaeA*, *ler*, *hylB*, and *qseA*), but no modification in Shiga toxin production has been observed (98). CFS and lactic acid produced by *Lmb. reuteri* significantly inhibited uropathogenic *Ec* (UPEC), thus reducing the production of virulence factors involved in the adhesion process, such as adhesion outer membrane proteins A and X, urogenital tract adherence promoter factor type 1, and P fimbriae subunits (99). Furthermore, studies conducted on different cell lines have confirmed the anti-*Ec* activity of several *Lactobacillus* strains. The adhesion ability of two *Ec* strains on Hep-2 and T84 cells was reduced after pretreatment with *Lbc. acidophilus* and *Lcb. rhamnosus* (100). Whereas *Lactobacillus jensenii* and *Lactobacillus gasseri* inhibited adhesion of DAEC to HeLa cells, *Lmb. reuteri* also reduced *Ec* internalization in the same cell line (101). Also, *Lpb. plantarum* and *Lcb. rhamnosus* inhibited *Ec* adherence to HT-29 cells by increasing the expression of intestinal mucins MUC2 and MUC3 (102). Also, an interference of induced cell signaling against DAEC caused by *Lbc. acidophilus* abolished the structural and functional microvilli alteration in human enterocyte-like cells (103, 104). As also reported for *Slm*, CLA overproducer *Lcb. casei* strain altered biofilm formation and modified *Ec* adhesion and invasion in INT407 cells (68). The combination of *Lcb. rhamnosus* with oligosaccharides resulted in an effective antidiarrheal formulation, owing to the increased autoaggregation and coaggregation properties of this strain. The inhibition of adherence to HT-29 cells was maximal with a *Lcb. rhamnosus* and inulin combination and significantly decreased the production of cyclic AMP, cyclic GMP, and related toxins (105). In an *in vitro* EHEC infection model, *Lcb. rhamnosus*, *Lbc. gasseri*, *Lcb. casei*, and *Lpb. plantarum* have been studied on C2BBe1 human colon epithelial cells. Among the tested strains, live *Lcb. rhamnosus* cells significantly reduced pathogen internalization, whereas this effect has not been observed with dead *Lcb. rhamnosus* cells or conditioned medium, thus implying that lactobacilli modulate the intracellular mechanism responsible for EHEC internalization (106). Multiple lactobacilli were also effective in inhibiting the *Ec* quorum sensing system, such as *Llb. sakei* and *Lbc. acidophilus* cell extract, which

significantly inhibited AI-2-like activity without affecting EHEC growth. Moreover, *Lbc. acidophilus* cell extracts inhibited biofilm formation on abiotic surfaces and HT-29 cell adhesion and downregulated the expression of several virulence factors associated with AI-2-like activity, particularly proteins involved in sulphur metabolism and membrane-associated functions (107, 108). *In vivo* experiments have shown similar results, including a significant decrease in adhesion and improvements in the immune system of infected animals. In a murine model, *Lactobacillus kefiranofaciens* treatment prevented EHEC infection-induced symptoms, Shiga toxin penetration, bacterial translocation, renal and intestinal damage, and increased mucosal EHEC-specific IgA responses. Lactobacilli also had protective effects in Caco-2 cells, reducing cell death and epithelial integrity loss induced by the pathogen (109). The ability of *Ec* to adhere to pig intestine brush borders decreased in a dose-dependent manner after administration of recombinant engineered fimbriae-producing *Lbc. acidophilus* (110). In an *in vivo* trial, the ability of *Ec* to disrupt the intestinal barrier and increase permeability was significantly reduced by administering *Lpb. plantarum* to rats, indicating a beneficial effect on the intestinal tract (111). *Lcb. casei* Shirota treatment of *Ec* in a murine urinary tract infection model inhibited growth and reduced inflammatory responses (112). In addition, exopolysaccharides produced during fermentation demonstrated *in vivo* anti-*Ec* activity, as reuterin and levan from *Lmb. reuteri* contained in weanling pig feed that reduced the number of *Ec* and the amount of heat-stable enterotoxin in colonic digesta (113). In addition, *Lcb. casei* strains decreased virulence gene expression in EHEC, *Slm*T, and *Lm*, particularly downregulating the *Ec* eaeA, *Slm*T *nmpC*, and *Lm fbp* and *iap* genes (114). Also, pretreatment of Caco-2 cells with live and heat-killed *Lbc. acidophilus* dose-dependently inhibited the adhesion and invasive properties of EPEC, *Lm*, *Slm*T, and *Yersinia pseudotuberculosis* (115, 116). Another study investigating the effect of pre-treatment of Caco-2 and HT-29 cells with lactobacilli reported that one *Lvb. brevis*, two *Lpb. plantarum*, and two *Lcb. paracasei* strains inhibited EPEC and *Slm*E adhesion to both cell lines (117).

1.3.5 *Clostridium* spp.

Hospital-acquired infections have severe consequences for already debilitated patients, and several studies have shown the effectiveness of lactobacilli in preventing the onset of such complications, as in the case of *Clostridium difficile* (*Cd*). This nosocomial bacterium infects the human gastrointestinal tract (118) and is characterized by two major virulence factors: enterotoxin A, expressed by the *tdcA* gene and causes diarrhea and intestinal mucosa damage, and toxin B, expressed by the *tcdB* gene and has strong cytotoxic effects (119, 120). Another important virulence factor is the *txeR* gene, which encodes a sigma factor that directs RNA polymerase to recognize the promoters of the *tcdA* and *tcdB* genes (121). Several lactobacilli have inhibitory effects on *Cd* virulence factors (Table S5 in the

supplemental material), particularly on the production of toxins, as demonstrated by various *in vitro* studies. Coculture of lactobacilli with *Cd* on Vero cells significantly decreased TcdA and TcdB toxins in spent supernatants and increased their intracellular concentrations, thereby suggesting a possible antagonistic mechanism that could reduce the synthesis and/or secretion of toxins (122). S-layer proteins extracted from *Lentilactobacillus kefiri* strains inhibited the damage caused by *Cd*-spent culture supernatants in Vero cells, and this activity was higher in aggregating strains than in nonaggregating strains, thus indicating a direct interaction between S-layer proteins and clostridial toxins. The same results were not obtained with live *Lbc. kefiri* cells, thereby indicating a different interaction between the soluble S-layer proteins and those located on the surface of the bacterium (123). *Lbc. acidophilus* CFS significantly reduced the cytotoxic and cytopathic effects of a hypervirulent *Cd* strain culture filtrate on human epithelial cells by decreasing pathogen attachment on HT-29 and Caco-2 cells (124). Inhibition of *Cd* virulence factors has also been observed *in vivo*. The administration of *Lbc. acidophilus* in *Cd*-inoculated mice altered QS molecule production, lowering the transcriptional levels of *luxS*, *tcdA*, *tcdB*, and *txeR* genes and increasing mouse survival ratios by as much as 80% (125). Furthermore, the administration of *Lmb. reuteri* significantly decreased *Cd* colonization and concentrations of toxins in the cecum and decreased the numbers of rotavirus, a human virus that causes gastroenteritis in infants and children, after both pretreatment and coincubation of the pathogen and the probiotic with HT-29 cells (126). In a protection model, an engineered Lactobacillus strain expressing TcdB-neutralizing antibody fragments delayed the death of infected hamsters (127), whereas in mice, an engineered *Lcb. casei* expressing Clostridium perfringens alpha-toxin toxoid induced the production of antibodies capable of neutralizing C. perfringens alpha-toxin and increasing levels of cytokines and interferon-g in the serum and spleen lymphocytes (128).

1.3.6 *Staphylococcus aureus*

Staphylococcus aureus (*Sa*) is an opportunistic pathogen accounting for 76% of all skin and soft tissue infections in humans (129) due to the expression of several virulence factors regulated by the *agr* QS system and the sae gene (130, 131). *Sa* produces a variety of toxins, including *sea* enterotoxins, which cause food poisoning (132), toxic shock syndrome toxin 1 (TSST-1) expressed by the *tst* gene, a superantigen that causes multiple organ dysfunctions and is associated with a high mortality rate (133), and *Staphylococcus* superantigen-like protein 1 (SSL-1), which inhibits the activity of matrix metalloproteases (134). The ability to evade the host immune system is promoted by the production of protein A (*spa*), a surface protein that prevents phagocytosis, and immunoglobulin- binding protein (*sbi*), which binds IgG and is involved in blood coagulation (135). Furthermore, the *mecA* gene

confers methicillin resistance to *Sa* (136), and the expression of the *ica* operon promotes biofilm formation (137). Several studies demonstrated that lactobacilli can effectively counteract the virulence factors of this pathogen (Table S6 in the supplemental material). Either cocultivation or CFS from different lactobacilli strains inhibited *Sa* biofilm formation, as in the case of the cocultivation with *Lcb. rhamnosus* (138) and acid CFS from *Lbc. acidophilus* that also inhibited lipase from biofilm and planktonic cells with a significant effect on methicillin-resistant *Sa* (139). In a study conducted on CFS produced by *Lpb. plantarum*, inhibition of the growth of *Sa* was observed, whereas CFS produced by *Lmb. fermentum* inhibited the expression of the *icaA* and *icaR* operons, thus limiting biofilm formation (140). CFS obtained from *Lpb. plantarum*, *Lmb. fermentum*, and *Lmb. reuteri* strains dependently decreased the expression of the *sea*, *sae*, *agrA*, *tst*, *spa*, and *spi* genes (141), and, in particular, the production of SSL-1 was significantly reduced when *Sa* was grown in *Lmb. reuteri* supernatant (142). Furthermore, *Lbc. acidophilus* and *Lmb. fermentum* have demonstrated a significant reduction of *Sa* adherence even on abiotic surfaces, most notably catheters and surgical implants (143, 144), thus suggesting a potential for the application of lactobacilli in the medical field to prevent the spread of nosocomial infections. The inhibitory effect of lactobacilli on *Sa* has also been confirmed *in vitro*. For example, *Lactobacillus crispatus* and *Lbc jensenii* coaggregated with *Sa*, preventing pathogen adhesion to vaginal cells (145), whereas live *Lcb. casei* cells affected *Sa* internalization, and both live and heat-killed *Lcb. casei* cells reduced *Sa* adhesion in bMEC cells (146). Depending on their growth phase, concentration, competition, and the presence of surface layer proteins, *Lgb. salivarius* and *Lpb. plantarum* significantly inhibited *Sa* adherence to Caco-2 cells (147).

1.3.7 *Helicobacter* spp.

Helicobacter is an important genus involved in food-borne illness. The clinical manifestations are determined by the genetics and behaviours of the human hosts (i.e., diet or smoking status) as well as bacterial virulence. *cagA* and *vacA* cytotoxin-associated genes are important in this regard; *cagA* alters intracellular signal transduction, and *vacA* induces the fusion between endosomes and lysosomes (148). Another important virulence factor is the production of flagellin, which is induced by the expression of *flaA* and *flaB* genes and provides the motility necessary for stomach colonization (149). Several studies have provided clear evidence that lactobacilli and their metabolites could decrease virulence factors of this species (Table S7 in the supplemental material). For example, the compounds produced by a *Lcb. casei* strain reduced the expression of genes codifying for flagellins in *Helicobacter* pylori (*flaA* and *flaB*) and *Slm*T (*flaC*), decreasing the motility and related internalization abilities (150). Similar results were obtained from a *Lmb. reuteri* strain, which

significantly reduced the expression of *flaA* and *vacA* genes (151), whereas *Lactiplantibacillus paraplantarum* CFS reduced the adherence of *H. pylori* on AGS cells (152). Pretreatment with live and UV-killed *Lgb. salivarius* strains promoted the modification of the interleukin and chemokine response in the same cell line, in addition to downregulating 8 of 12 genes belonging to the *H. pylori Cag* pathogenicity island. This immunomodulatory effect was not dependent on adhesion or bacteriocin production, but after *Lgb. salivarius* exposure, CagA protein accumulated inside *H. pylori* cells, probably because of the loss of CagA secretion functionality (153). *In vivo* tests on *Helicobacter hepaticus*-stimulated macrophages from IL-10-deficient mice have been performed to investigate TNF-a-inhibitory *Lmb. reuteri* and *Lcb. paracasei*. These lactobacilli effectively decreased intestinal inflammation by lowering the levels of the proinflammatory colonic cytokines TNF-a and IL-12 but had no effects on *H. hepaticus* vitality (154). *Lbc. acidophilus* eradicated *H. pylori* from colonized children in 6.5% of subjects, while no spontaneous clearance was observed in untreated children, demonstrating the efficacy of lactobacilli administration in humans (155).

1.3.8 *Pseudomonas* spp., *Streptococcus* spp., and *Klebsiella* spp.

Biofilms are microorganism aggregations within an extracellular matrix composed of proteins, exopolysaccharides, water, nutrients (such as polysaccharides and amino acids), and ions. The ability to form biofilms is an important common property that increases pathogen virulence, conferring adhesiveness and resistance to the host immune system and antibiotics (156). Biofilm formation is a characteristic trait of *Pseudomonas* spp., *Streptococcus* spp., and *Klebsiella* spp., all of which can establish ecological niches in which they replicate and become infectious to humans. Also in this case, lactobacilli and their metabolites have proven to be effective in inhibiting specific virulence factors of these pathogens (Table S8 in the supplemental material). *Pseudomonas aeruginosa*, one of the most common pathogens in the hospital setting, owes its pathogenicity to various virulence factors (besides biofilm formation), such as the secretion of toxins (157) and the presence of flagella and pili (158). *P. aeruginosa* biofilm formation and elastase production were effectively inhibited by *Lmb. fermentum*, *Lbc. zeae*, and *Lcb. paracasei* (159), whereas *Apilactobacillus kunkeei* exhibited *in vitro* antibiofilm properties and attenuated *P. aeruginosa* infection in a *G. mellonella* model (160). Other *in vivo* tests were performed to evaluate the effects of *Lpb. plantarum* on *P. aeruginosa* acyl-homoserine-lactones, elastases, and biofilm virulence factors. In a burned mouse model, lactobacilli inhibited *P. aeruginosa* colonization, thus improving tissue repair and enhancing pathogen phagocytosis (161). Crude extract from *Companilactobacillus crustorum* degraded N-homoserine lactone and significantly enhanced biofilm sensitivity to azithromycin, thereby inhibiting biofilm formation and reducing the thickness of already formed biofilms. Real-time quantitative PCR (RT-

qPCR) analysis revealed downregulation of *lasI/R* and *rhlI/R* QS virulence genes as well as inhibition of chitinase, protease, rhamnolipid, alginate, pyocyanin, and exopolysaccharide synthesis (162). *Klebsiella pneumoniae*, a pathogenic bacterium associated with urinary infections that occur primarily in hospitalized patients and are frequently connected with the use of medical devices, is another microorganism whose pathogenicity relies on the ability to form biofilms (163). In this regard, *Lmb. fermentum* cells and their acid supernatants exerted antibiofilm properties against *K. pneumoniae* on catheters (164). In addition, *Lbc. acidophilus* and *Lmb. fermentum* or their supernatants hindered pathogen spread within biofilms, since no *K. pneumoniae* live cells were found after treatment (165). *Streptococcus mutans* is the main etiological agent of human dental caries, owing to its virulence factors such as the aforementioned ability to form biofilms (166) as well as glucosyltransferases encoded by *gtf* and *tft* genes, which enable the production of exopolysaccharides and thus the formation of plaque (167), and fructosyltransferase (*ftf*), which is essential in adhesion (168). Different lactobacilli produce biosurfactants that downregulate the expression of *S. mutans* biofilm-forming genes, for example, *Lmb. fermentum* and *Lbc. acidophilus*, which reduced *gtfB* and *gtfC* gene expression modifying the surface and adhesion properties of the pathogen (169, 170), *Lmb. reuteri*, which reduced *gftB*, *gtfC*, and *fft* gene expression (168), and *Lbc. acidophilus*, which downregulated *gtf* and *luxS* (171). Similar results were obtained with the coculture of *S. mutans* with *Lcb. casei*, which downregulated *luxS* and *gftB*, *spaP*, and *gbpB* adhesion genes (172). Likewise, *Lcb. casei*, *Lmb. reuteri*, *Lpb. plantarum*, *Lgb. salivarius*, *Lcb. rhamnosus*, and *Lmb. reuteri* decreased biofilm formation and downregulated the *gtf* genes, significantly decreasing bacterial attachment to surfaces (173–175). Lactobacilli were also effective against *Streptococcus pyogenes*, a pathogen that affects humans exclusively and causes a variety of disorders ranging from asymptomatic transport to mild and superficial infections of the skin and mucous membranes to systemic diseases (176). Its virulence depends on the production of toxins, in particular streptolysin S encoded by the *sag* operon, which causes erythrocytes lysis (177). The combination of *Lcb. rhamnosus* and *Lmb. reuteri* and their spent media were the most effective in reducing *S. pyogenes* adherence in FaDu and Detroit 562 host cells, inhibiting hemolytic activity through the downregulation of *sag* operon expression with a consequent decrease in streptolysin S production (178). In addition, a *Lpb. plantarum* strain decreased the levels of IL-17 and IL-23 in Hep-2 and A549 cells exposed to *S. pyogenes* by inducing the Tolllike receptor 2 (TLR2)/TLR4 surface receptors involved in the immune response (179).

1.3.9 Urogenital-correlated pathogens

Urogenital tract infections are major causes of disease in women. Several pathogenic species, including *Candida albicans*, *Chlamydia trachomatis*, *Ec*, *Gardnerella vaginalis*, *Neisseria*

gonorrhoeae, *Prevotella bivia*, *Streptococcus agalactiae*, and *Trichomonas vaginalis*, are involved in the onset of disorders that, if untreated, can cause serious irreversible complications (180). In healthy individuals, the vaginal microbiota is dominated by lactobacilli (181), which protect against infections by inhibiting pathogen colonization via several mechanisms (Table S9 in the supplemental material), such as increasing microbiota adhesion through the production of biosurfactants, competition for host cell receptors, or direct killing through the production of hydrogen peroxide and bacteriocins (182). Inhibition of pathogen adhesion has been observed both in cell lines and on abiotic surfaces. *Lbc. acidophilus*, *Lbc. gasseri*, and *Lbc. jensenii* isolated from the human vagina were able to autoaggregate and strongly adhere to vaginal cell surfaces (183), whereas *Lpb. plantarum* coaggregated with pathogens such as *S. agalactiae*, *G. vaginalis*, and *Ec* (184). Moreover, a *Lbc. acidophilus* strain was able to inhibit *Staphylococcus epidermidis* and UPEC attachment on abiotic surfaces (185). Other urogenital tract pathogens include *Trichomonas vaginalis*, which causes trichomoniasis, *Neisseria gonorrhoeae*, which causes gonorrhea, and *Gardnerella vaginalis*, which is responsible for the initiation of bacterial vaginosis due to its ability to form biofilm. The most important virulence factor of *T. vaginalis* and *N. gonorrhoeae* is vaginal cell adhesion ability (186, 187), whereas *G. vaginalis* produces vaginolysin (*vly*), which inhibits the immune response, and sialidase (*sld*), an enzyme that releases salicylic acid, which improves adherence to cells and surfaces. Lactobacilli isolated from the human vagina showed significant inhibitory activities toward *T. vaginalis*, *N. gonorrhoeae*, and *G. vaginalis*. In particular, pretreatment with *Lbc. crispatus* competitively excluded *G. vaginalis* adhesion to HeLa cells, reducing the expression of *vly* and *sld* virulence genes (188), whereas *Lbc. gasseri* and *Lbc. jensenii* inhibited adhesion of *T. vaginalis* and *N. gonorrhoeae* to VEC and Hec-1-B cell lines, respectively (189, 190). Furthermore, a recombinant *Lbc. jensenii* secreting two domain CD4 proteins prevented the entrance of human immunodeficiency virus (HIV) into HeLa cells (191). Different trials observed the ability of *Lbc. gasseri*, *Lbc. crispatus*, and *Lbc. helveticus* to counteract vaginal-associated pathogens, specifically protecting cervix epithelial cells against the effects of *P. bivia*, toxin-producing *G. vaginalis*, and UPEC, inhibiting their adhesion to HeLa cells (192, 193). Similar results were obtained from *Lbc. helveticus*, which was able to inhibit the adhesion of *G. vaginalis* and UPEC to HeLa cells and internalization of UPEC and *Slm*T on HeLa and Caco2 cells, respectively (194). *Candida albicans* is an opportunistic pathogenic yeast that resides in the oral cavity and gastrointestinal and urogenital tracts and is responsible for oral and vulvovaginal candidiasis. Its pathogenicity arises from multiple factors, including adherence promoted by various types of adhesins (*Als3* and *Hwp1*), biofilm formation (*Ece1*, *Als3*, *Bcr1*, *Efg1*, *Tec1*, and *Cph1*), resistance to drugs, and the immune system through overexpression of Cdr1, Cdr2, and Mrd1 proteins (195), yeast-to-hyphal morphogenesis (*Ece1*), and

hydrolytic enzymes (*Saps*) (196). Probiotic lactobacilli are effectively used in medical treatments to limit the spread of *C. albicans* by maintaining the balance of microbiota and producing inhibitory substances active against the pathogen (197–199). Lactobacilli isolated from women produced biosurfactants that significantly reduced *C. albicans* adhesion and prevented the formation of biofilms, and maximal results were obtained with *Lbc. gasseri*, *Lmb. reuteri*, *Lbc. acidophilus*, and *Lcb. paracasei* (200). Similar effects were obtained by coinoculating *Lpb. plantarum*, *Lmb. fermentum*, *Lbc. gasseri*, and *Lmb. reuteri* with *C. albicans*. Their autoaggregative properties, enhanced by low pH values and biofilm-forming ability, resulted in vaginal tract colonization, whereas coaggregation with *C. albicans* prevented yeast adhesion (201). *Lbc. gasseri* and *Lbc. crispatus* CFS coincubation with *C. albicans* significantly reduced the expression of *Hwp1* and *Ece1*, *Als3*, *Bcr1*, *Efg1*, *Tec1*, and *Cph1* genes, lowering biofilm formation, whereas CFS from *Lbc. crispatus* inhibited *C. albicans* adhesion to HeLa cells (202). Another important mechanism of virulence inhibition is the modification of the hyphal structure. Several studies found that *Lcb. rhamnosus* reduced hyphal elongation (203), and *Lcb. rhamnosus*, *Lcb. paracasei*, and *Lcb. casei* were effective against *C. albicans* hyphal morphogenesis because they expressed the *MspI* gene, encoding a major peptidoglycan hydrolase that hydrolyzes chitin (204). Proteinase and hemolysin activities were reduced in *C. albicans* grown with *Lcb. rhamnosus*, with alterations to antifungal susceptibility (205). In addition, *Lcb. rhamnosus* affected adhesion, invasion, and hyphal extension, preventing oral epithelial tissue damage. This effect was correlated with glucose depletion and repression of ergosterol synthesis (206). Several lactobacilli had different effects on *C. albicans*-induced interleukin in VK2/E6E7 cells: for example, *Lcb. rhamnosus* alone or in combination with *Lmb. reuteri* inhibited the increase in IL-1α and IL-8, whereas their supernatants increased IL-8 and IP-10 levels (207). In addition, *Lbc. crispatus* lowered *C. albicans* adhesion to VK2/E6E7 cells, thus upregulating IL-2, IL-6, and IL-17 while downregulating IL-8 (208), and to HeLa cells, lowering IL-8 and increasing β-defensin 2 and 3 (209). In the same cell line, a reduction in adhesion was attributed to antifungal activity arising from the inhibition of histone deacetylase by *Lbc. crispatus*, *Limosilactobacillus vaginalis*, and *Lbc. gasseri* (210). Several studies have investigated the effects of lactobacilli on gene expression of this pathogen. An extract from a *Lactobacillus* species strain, owing to high levels of oleic and myristic acid, affected *C. albicans* virulence (hyphal formation, proteinase, and phospholipase secretion), thus decreasing also *Hwp1*, *Plb2*, and *Sap1* virulence genes expression (211). Moreover, CFSs of *Lbc. crispatus*, *Lbc. gasseri*, *Lbc. acidophilus*, and *Lbc. jensenii* effectively decreased the yeast-to-hyphal transition and the expression of hyphae-specific genes *Als3*, *Hwp1*, and *Ece1*, whereas *Nrg1*, a negative transcriptional regulator, was upregulated (212). *Lcb. rhamnosus* and its supernatant reduced *C. albicans* filamentation and biofilm formation *in vitro*, altering the

expression of *Bcr1*, *Hwp1*, and *Als3* adhesion genes and *Cph1* transcriptional regulatory genes. The same strain was tested on *G. mellonella* infected with *C. albicans*, and this treatment increased larval survival up to 80% (213). *Lcb. paracasei*, *Lmb. fermentum*, and *Lcb. rhamnosus* also attenuated candidiasis in *G. mellonella* by increasing hemocyte quantity, upregulating galiomicin and gallerymicin antifungal peptide genes, slowing hyphal formation, and lowering biofilm development by downregulating the *Als3*, *Hwp1*, *Efg1*, and *Cph1* genes (214). In other studies, *Lbc. acidophilus* and its filtrate inhibited *C. albicans* filamentation and biofilm formation, increasing the *G. mellonella* survival rate (215).

1.3.10 Other pathogens

Multiple studies have been conducted on other pathogens and have shown encouraging results (Table S10 in the supplemental material). The modulating effect of lactobacilli on the immune system had positive effects in both mice inoculated with *Yersinia enterocolitica* and children infected with *Enterococcus faecalis*. In the first case, *Lpb. plantarum* had an immunomodulatory effect on infected BALB/c mice, resulting in a decrease in the anti-inflammatory cytokine IL-10 and an increase in IgA production (216). The administration of *Lcb. rhamnosus* to children colonized with vancomycin-resistant *En. faecalis* led to immune system modulation, preventing the onset of infection (217). *Lpb. plantarum* also increased the virulence of *Serratia marcescens*, which causes hospital-acquired infections and whose antibiotic resistance poses a severe risk to patients, and of *Bacillus cereus*, which causes food poisoning. In relation to inoculum concentration and temperature, *Lpb. plantarum* reduced the hemolytic activity and protease and lecithinase expression of *B. cereus* (218), whereas CFS from *Lbc. acidophilus* and *Lpb. plantarum* affected the resistance of *Se. marcescens* to ceftriaxone and completely inhibited swarming motility (219). In addition, the CFS of *Lgb. salivarius* and *Lbc. gasseri* significantly reduced the virulence gene expression of *Aggregatibacter actinomycetemcomitans*, an oral pathogen that causes localized periodontitis by producing leukotoxins (LtzA) and cytolethal distending toxin (CdtB) (220).

1.4 Conclusions

Despite the development of various effective therapies, bacterial infections continue to pose a major threat to public health. In this regard, as described herein, lactobacilli capable of counteracting the virulence abilities of pathogenic microorganisms could be used to support existing treatments. Some of these mechanisms include the reduction of the adhesive and invasive properties, the ability to self-aggregate and coaggregate with the pathogens, direct downregulation of virulence genes, and the production of metabolites with specific activities that can affect and modulate the host immune

response. In addition, their presence has a bioprotective effect on both abiotic surfaces and cellular tissues. Lactobacilli, through competition for substrate and their steric hindrance, can inhibit pathogen activity and reduce their ability to adhere to epithelial cells, hence preventing the onset of diseases. Although from review of the literature, many authors have demonstrated the ability to reduce virulence factors in pathogens by lactobacilli (our sincere apologies go to colleagues whose work was involuntarily not cited); however, there are still few studies conducted directly on humans validating all these capabilities observed in *in vitro* and *in vivo* tests on animals. Further research on this topic would thus help understand and advance the real applications of this microbial group to counteract pathogen virulence. Lactobacilli, which have always been used by mankind and have a long history of safe use by humans in food preservation and processing, are currently also used as probiotics thanks to their proven beneficial properties. In addition to this, current whole-genome sequencing techniques provide additional assurance of safety, as evidenced by the recent EFSA statement, which recommends genetic characterization of all microbial strains before their use in food applications (221). Knowledge of the whole genome enables the identification of all potential risk factors present in lactobacilli (222), thus increasing the safety of use even in debilitated patients in hospital settings, where complete safety of the bacterial strains used must be ensured. In fact, beyond the current use as probiotics to alleviate the adverse effects of antibiotic therapies, lactobacilli could be used also as adjuvants for antibiotics, owing to their ability to counteract pathogens and their virulence properties. Infectious disease prevention is a fundamental achievement to limit the widespread use of drugs to strictly necessary cases, thus hindering the spread of antibiotic resistance. This issue has made treatment of infection more difficult in recent years; therefore, identifying alternative treatments is increasingly important to decrease the use of antibiotics while also improving host health. Given that the average age of the world population is rising, the consequences of demographic aging are expected to have severe repercussions on numerous social dynamics in the future, including an increase in the cost of public health. To reduce the number of hospitalizations and consequently the costs of health care, the condition of older and fragile people must be improved. The identification and study of strains with probiotic and antivirulence activity against pathogens may lead to the development of therapies that can be combined with current antibiotic treatments, thus reducing their adverse effects on patients while increasing their effectiveness. Furthermore, consistent intake of strains capable of reducing the likelihood of pathological manifestations in hosts, such as through the consumption of food formulations, could also be used to prevent infections, thereby reducing antibiotic use.

1.5 References

1. Zheng J, Wittouck S, Salvetti E, Franz CMAP, Harris HMB, Mattarelli P, O'toole PW, Pot B, Vandamme P, Walter J, Watanabe K, Wuyts S, Felis GE, Gänzle MG, Lebeer S. 2020. 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 70:2782–2858.
2. Plavec T V., Berlec A. 2020. Safety Aspects of Genetically Modified Lactic Acid Bacteria. Microorganisms 8.
3. EFSA BIOHAZ Panel (Panel on Biological Hazards). 2016. Update of the list of QPS-recommended biological agents intentionally added to food or feed as notified to EFSA 4: suitability of taxonomic units notified to EFSA until March 2016. EFSA J 14.
4. FAO/WHO. 2001. Joint FAO/WHO Expert Consultation on Evaluation of Health and Nutritional Properties of Probiotics in Food including Powder Milk with Live Lactic Acid BacteriaFao Food and Nutrition Paper.
5. Hill C, Guarner F, Reid G, Gibson GR, Merenstein DJ, Pot B, Morelli L, Canani RB, Flint HJ, Salminen S, Calder PC, Sanders ME. 2014. Expert consensus document: The international scientific association for probiotics and prebiotics consensus statement on the scope and appropriate use of the term probiotic. Nat Rev Gastroenterol Hepatol 11:506–514.
6. Salminen S, Collado MC, Endo A, Hill C, Lebeer S, Quigley EMM, Sanders ME, Shamir R, Swann JR, Szajewska H, Vinderola G. 2021. The International Scientific Association of Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of postbiotics. Nat Rev Gastroenterol Hepatol.
7. Liévin-Le Moal V, Servin AL. 2014. Anti-infective activities of Lactobacillus strains in the human intestinal microbiota: From probiotics to gastrointestinal anti- infectious biotherapeutic agents. Clin Microbiol Rev 27:167–199.
8. Di Cerbo A, Palmieri B, Aponte M, Cesar Morales-Medina J, Iannitti T. 2016. Mechanisms and therapeutic effectiveness of lactobacilli. J Clin Pathol 69:187–203.
9. Gao K, Mu CL, Farzi A, Zhu WY. 2020. Tryptophan Metabolism: A Link between the Gut Microbiota and Brain. Adv Nutr 11:709–723.
10. Turroni F, Ventura M, Buttó LF, Duranti S, O'Toole PW, Motherway MOC, Van Sinderen D. 2014. Molecular dialogue between the human gut microbiota and the host: A Lactobacillus and Bifidobacterium perspective. Cell Mol Life Sci 71:183–203.
11. Slover CM. 2008. Lactobacillus: a Review. Clin Microbiol Newsl 30:23–27.
12. de Melo Pereira GV, de Oliveira Coelho B, Magalhães Júnior AI, Thomaz-Soccol V, Soccol

CR. 2018. How to select a probiotic? A review and update of methods and criteria. Biotechnol Adv 36:2060–2076.

13. Masood MI, Qadir MI, Shirazi JH, Khan IU. 2011. Beneficial effects of lactic acid bacteria on human beings. Crit Rev Microbiol 37:91–98.
14. Ohashi Y, Ushida K. 2009. Health-beneficial effects of probiotics: Its mode of action. Anim Sci J 80:361–371.
15. Hilmi HTA. 2010. Lactic Acid Bacteria and Their Antimicrobial Peptides : Induction , Detection , Partial Characterization , and Their Potential Applications.
16. Collado MC, Surono I, Meriluoto J, Salminen S. 2007. Indigenous Dadih Lactic AcidBacteria: Cell-Surface Propertiesand Interactions with Pathogens. J Food Sci 72.
17. Waters CM, Bassler BL. 2005. Quorum sensing: Cell-to-cell communication in bacteria. Annu Rev Cell Dev Biol.
18. Vendeville A, Winzer klaus, Heurlier K, Tang CM, Hrdie KR. 2005. Making “sense” of metabolism: autoinducer-2, LUXS and pathogenic bacteria. Nat Rev Microbiol 3:383–396.
19. Verderosa AD, Totsika M, Fairfull-Smith KE. 2019. Bacterial Biofilm Eradication Agents: A Current Review. Front Chem 7.
20. Carpentier B, Chassaing D. 2004. Interactions in biofilms between Listeria monocytogenes and resident microorganisms from food industry premises. Int J Food Microbiol 97:111–122.
21. Dicks LMT, Grobbelaar MJ. 2021. Double-barrel shotgun: Probiotic lactic acid bacteria with antiviral properties modified to serve as vaccines. Microorganisms. MDPI AG.
22. WHO. 2017. World Health Statistics 2017 148:148–162.
23. Hutchings MI, Truman AW, Wilkinson B. 2019. Antibiotics: past, present and future. Curr Opin Microbiol 52:72–80.
24. Rello J, Parisella FR, Perez A. 2019. Alternatives to antibiotics in an era of difficult-to-treat resistance: new insights. Expert Rev Clin Pharmacol. Taylor and Francis Ltd.
25. Syed B, Wein S, Ruangapaint Y. 2020. The Efficacy of Synbiotic Application in Broiler Chicken Diets , 10:469–479.
26. Disson O, Moura A, Lecuit M. 2021. Making Sense of the Biodiversity and Virulence of Listeria monocytogenes. Trends Microbiol. Elsevier Ltd.
27. Camejo A, Carvalho F, Reis O, Leitão E, Sousa S, Cabanes D. 2011. The arsenal of virulence factors deployed by Listeria monocytogenes to promote its cell infection cycle. Virulence 2:379–394.
28. Piercey MJ, Hingston PA, Truelstrup Hansen L. 2016. Genes involved in Listeria monocytogenes biofilm formation at a simulated food processing plant temperature of 15 °C.

Int J Food Microbiol 223:63–74.
29. Quereda JJ, Andersson C, Cossart P, Johansson J, Pizarro-Ceredà J. 2018. Role in virulence of phospholipases, listeriolysin O and listeriolysin S from epidemic Listeria monocytogenes using chicken embryo infection model. Vet Res 49.
30. Jacquet C, Gouin E, Jeannel D, Cossart P, Rocourt J. 2002. Expression of ActA, Ami, InlB, and Listeriolysin O in Listeria monocytogenes of human and food origin. Appl Environ Microbiol 68:616–622.
31. Sibanda T, Buys EM. 2022. Listeria monocytogenes Pathogenesis : The Role of Stress Adaptation.
32. Corr S, Hill C, Gahan CGM. 2006. An in vitro cell-culture model demonstrates internalin- and hemolysin-independent translocation of Listeria monocytogenes across M cells. Microb Pathog 41:241–250.
33. Dutra V, Silva AC, Cabrita P, Peres C, Malcata X, Brito L. 2016. Lactobacillus plantarum LB95 impairs the virulence potential of Gram-positive and Gram-negative food-borne pathogens in HT-29 and vero cell cultures. J Med Microbiol 65:28–35.
34. Iglesias MB, Viñas I, Colás-Medà P, Collazo C, Serrano JCE, Abadias M. 2017. Adhesion and invasion of Listeria monocytogenes and interaction with Lactobacillus rhamnosus GG after habituation on fresh-cut pear. J Funct Foods 34:453–460.
35. Dong Q, Zhang W, Guo L, Niu H, Liu Q, Wang X. 2020. Influence of Lactobacillus plantarum individually and in combination with low O2-MAP on the pathogenic potential of Listeria monocytogenes in cabbage. Food Control 107:106765.
36. Upadhyay A, Upadhyaya I, Mooyottu S, Venkitanarayanan K. 2016. Eugenol in combination with lactic acid bacteria attenuates Listeria monocytogenes virulence in vitro and in invertebrate model Galleria mellonella. J Med Microbiol 65:443–455.
37. Koo OK, Amalaradjou MAR, Bhunia AK. 2012. Recombinant probiotic expressing Listeria adhesion protein attenuates Listeria monocytogenes virulence in vitro. PLoS One 7.
38. Mathipa MG, Bhunia AK, Thantsha MS. 2019. Internalin AB-expressing recombinant Lactobacillus casei protects Caco-2 cells from Listeria monocytogenes-induced damages under simulated intestinal conditions. PLoS One 14:e0220321.
39. Archambaud C, Nahori MA, Soubigou G, Bećavin C, Laval L, Lechat P, Smokvina T, Langella P, Lecuit M, Cossart P. 2012. Impact of lactobacilli on orally acquired listeriosis. Proc Natl Acad Sci U S A 109:16684–16689.
40. Bambirra FHS, Lima KGC, Franco BDGM, Cara DC, Nardi RMD, Barbosa FHF, Nicoli JR. 2007. Protective effect of Lactobacillus sakei 2a against experimental challenge with Listeria

monocytogenes in gnotobiotic mice. Lett Appl Microbiol 45:663–667.

41. Riaz A, Noureen S, Liqat I, Arshad M, Arshad N. 2019. Antilisterial efficacy of Lactobacillus brevis MF179529 from cow: An in vivo evidence. BMC Complement Altern Med 19:1–9.
42. Deng Q, Shi H, Luo Y, Zhao H, Liu N. 2020. Effect of dietary Lactobacilli mixture on Listeria monocytogenes infection and virulence property in broilers. Poult Sci 99:3655–3662.
43. Alves VF, Lavrador MAS, De Martinis ECP. 2003. Bacteriocin exposure and food ingredients influence on growth and virulence of Listeria monocytogenes in a model meat gravy system. J Food Saf 23:201–217.
44. Martinez RCR, De Martinis ECP. 2005. Evaluation of bacteriocin-producing Lactobacillus sakei 1 against Listeria monocytogenes 1/2a growth and haemolytic activity. Brazilian J Microbiol 36:83–87.
45. Winkelströter LK, Gomes BC, Thomaz MRS, Souza VM, De Martinis ECP. 2011. Lactobacillus sakei 1 and its bacteriocin influence adhesion of Listeria monocytogenes on stainless steel surface. Food Control 22:1404–1407.
46. Martinez RCR, De Martinis ECP. 2005. Antilisterial activity of a crude preparation of Lactobacillus sakei 1 bacteriocin and its lack of influence on Listeria monocytogenes haemolytic activity. Food Control 16:429–433.
47. Andino A, Hanning I. 2015. Salmonella enterica: Survival, colonization, and virulence differences among serovars. Sci World J. Hindawi Publishing Corporation.
48. Lucas RL, Lee CA. 2001. Roles of hilC and hilD in regulation of hilA expression in Salmonella enterica serovar typhimurium. J Bacteriol 183:2733–2745.
49. Bergeron JRC, Brockerman JA, Vuckovic M, Deng W, Okon M, Finlay BB, McIntosh LP, Strynadka NCJ. 2018. Characterization of the two conformations adopted by the T3SS inner-membrane protein PrgK. Protein Sci 27:1680–1691.
50. Muyyarikkandy MS, Amalaradjou MA. 2017. Lactobacillus bulgaricus, lactobacillus rhamnosus and lactobacillus paracasei attenuate salmonella enteritidis, salmonella heidelberg and salmonella typhimurium colonization and virulence gene expression in vitro. Int J Mol Sci 18.
51. Guiney DG, Fierer J. 2011. The role of the spv genes in Salmonella pathogenesis. Front Microbiol 2.
52. Pati NB, Vishwakarma V, Jaiswal S, Periaswamy B, Hardt WD, Suar M. 2013. Deletion of invH gene in Salmonella enterica serovar Typhimurium limits the secretion of Sip effector proteins. Microbes Infect 15:66–73.
53. Lou L, Zhang P, Piao R, Wang Y. 2019. Salmonella Pathogenicity Island 1 (SPI-1) and Its

Complex Regulatory Network. Front Cell Infect Microbiol. Frontiers Media S.A.

54. Jiang L, Wang P, Song X, Zhang H, Ma S, Wang J, Li W, Lv R, Liu X, Ma S, Yan J, Zhou H, Huang D, Cheng Z, Yang C, Feng L, Wang L. 2021. Salmonella Typhimurium reprograms macrophage metabolism via T3SS effector SopE2 to promote intracellular replication and virulence. Nat Commun 12:1–18.
55. Haranga A, Miller S i. 2003. A Salmonella enterica Serovar Typhimurium translocated Leucine-Rich Repeat Effector Protein inhibits NF-kB-Dependent Gene Expression. Infect Immun 71:4052–4058.
56. Liao AP, Petrof EO, Kuppireddi S, Xia Y, Claus EC, Sun J. 2008. Salmonella type III effector AvrA stabilizes cell tight junctions to inhibit inflammation in intestinal epithelial cells. PLoS One 3.
57. Pérez-Morales D, Banda MM, Chau NYE, Salgado H, Martínez-Flores I, Ibarra JA, Ilyas B, Coombes BK, Bustamante VH. 2017. The transcriptional regulator SsrB is involved in a molecular switch controlling virulence lifestyles of Salmonella. PLoS Pathog 13.
58. SM A, M.O. A, A. A, Eldemerdash A. 2016. Molecular Studies on the prophylactic effect of probiotics on Salmonella typhimurium infected chicks. Benha Vet Med J 31:73–82.
59. Wang C, Wang J, Gong J, Yu H, Pacan JC, Niu Z, Si W, Sabour PM. 2011. Use of Caenorhabditis elegans for preselecting Lactobacillus isolates to control Salmonella Typhimurium. J Food Prot 74:86–93.
60. Yang X, Brisbin J, Yu H, Wang Q, Yin F, Zhang Y, Sabour P, Sharif S, Gong J. 2014. Selected lactic acid-producing bacterial isolates with the capacity to reduce salmonella translocation and virulence gene expression in chickens. PLoS One 9.
61. Andino A, Zhang N, Diaz-Sanchez S, Yard C, Pendleton S, Hanning I. 2014. Characterization and specificity of probiotics to prevent salmonella infection in mice. Funct Foods Heal Dis 4:370–380.
62. Song F, Liu J, Zhao W, Huang H, Hu D, Chen H, Zhang H, Chen W, Gu Z. 2020. Synergistic Effect of Eugenol and Probiotic Lactobacillus Plantarum Zs2058 Against Salmonella Infection in C57bl/6 Mice. Nutrients 12:1611.
63. Liévin-Le Moal V, Amsellem R, Servin AL. 2011. Impairment of Swimming Motility by Antidiarrheic Lactobacillus acidophilus Strain LB Retards Internalization of Salmonella enterica Serovar Typhimurium within Human Enterocyte-Like Cells. Antimicrob Agents Chemother 55:4810–4820.
64. Abdel-Daim A, Hassouna N, Hafez M, Ashor MSA, Aboulwafa MM. 2013. Antagonistic activity of lactobacillus isolates against salmonella typhi in vitro. Biomed Res Int 2013.

65. Burkholder KM, Fletcher DH, Gileau L, Kandolo A. 2019. Lactic acid bacteria decrease Salmonella enterica Javiana virulence and modulate host inflammation during infection of an intestinal epithelial cell line. Pathog Dis 77:2019.
66. Burkholder KM, Bhunia AK. 2009. Salmonella enterica serovar Typhimurium adhesion and cytotoxicity during epithelial cell stress is reduced by Lactobacillus rhamnosus GG. Gut Pathog 1:14.
67. Makras L, Triantafyllou V, Fayol-Messaoudi D, Adriany T, Zoumpopoulou G, Tsakalidou E, Servin A, De Vuyst L. 2006. Kinetic analysis of the antibacterial activity of probiotic lactobacilli towards Salmonella enterica serovar Typhimurium reveals a role for lactic acid and other inhibitory compounds. Res Microbiol 157:241–247.
68. Peng M, Tabashsum Z, Patel P, Bernhardt C, Biswas D. 2018. Linoleic acids overproducing lactobacillus casei limits growth, survival, and virulence of salmonella typhimurium and enterohaemorrhagic Escherichia coli. Front Microbiol 9:1–14.
69. Tabashsum Z, Peng M, Bernhardt C, Patel P, Carrion M, Rahaman SO, Biswas D. 2020. Limiting the pathogenesis of Salmonella Typhimurium with berry phenolic extracts and linoleic acid overproducing Lactobacillus casei. J Microbiol 58:489–498.
70. Grzesiuk E, Jankowska A, Laubitz D, Antushevich H, Zabielski R. 2008. Competition of Lactobacillus paracasei with Salmonella enterica for adhesion to Caco-2 cells. J Biomed Biotechnol 2008.
71. Coconnier-Polter MH, Liévin-Le Moal V, Servin AL. 2005. A Lactobacillus acidophilus strain of human gastrointestinal microbiota origin elicits killing of enterovirulent Salmonella enterica serovar typhimurium by triggering lethal bacterial membrane damage. Appl Environ Microbiol 71:6115–6120.
72. Durant JA, Corrier DE, Stanker LH, Ricke SC. 2000. Salmonella enteritidis hila gene fusion response after incubation in spent media from either S. Enteritidis or a poultry Lactobacillus strain. J Environ Sci Heal - Part B Pestic Food Contam Agric Wastes 35:599–610.
73. Hudault S, Liévin V, Bernet-Camard MF, Servin AL. 1997. Antagonistic activity exerted in vitro and in vivo by lactobacillus casei (strain GG) against Salmonella typhimurium C5 infection. Appl Environ Microbiol 63:513–518.
74. Craven SE, Williams DD. 1998. In vitro attachment of Salmonella typhimurium to chicken cecal mucus: Effect of cations and pretreatment with Lactobacillus spp. isolated from the intestinal tracts of chickens. J Food Prot 61:265–271.
75. Hu JL, Yu H, Kulkarni RR, Sharif S, Cui SW, Xie MY, Nie SP, Gong J. 2015. Modulation of cytokine gene expression by selected Lactobacillus isolates in the ileum, caecal tonsils and

spleen of Salmonella-challenged broilers. Avian Pathol 44:463–469.

76. Eryildiz C, Tabakcioglu K, Kuyucuklu G, Sakru N. 2020. Investigation of Antimicrobial Resistance and Virulence Genes of Campylobacter Isolates from Patients in Tertiary Hospital in Edirne, Turkey. Indian J Med Microbiol 38:157–161.
77. Mohan V. 2015. The role of probiotics in the inhibition of Campylobacter jejuni colonization and virulence attenuation. Eur J Clin Microbiol Infect Dis 34:1503–1513.
78. Alemka A, Clyne M, Shanahan F, Tompkins T, Corcionivoschi N, Bourke B. 2010. Probiotic colonization of the adherent mucus layer of HT29MTXE12 cells attenuates Campylobacter jejuni virulence properties. Infect Immun 78:2812–2822.
79. Taha-Abdelaziz K, Astill J, Kulkarni RR, Read LR, Najarian A, Farber JM, Sharif S. 2019. In vitro assessment of immunomodulatory and anti-Campylobacter activities of probiotic lactobacilli. Sci Rep 9:1–15.
80. Tabashsum Z, Peng M, Salaheen S, Comis C, Biswas D. 2018. Competitive elimination and virulence property alteration of Campylobacter jejuni by genetically engineered Lactobacillus casei. Food Control 85:283–291.
81. Mundi A, Delcenserie V, Amiri-Jami M, Moorhead S, Griffiths MW. 2013. Cell-free preparations of Lactobacillus acidophilus strain La-5 and Bifidobacterium longum strain NCC2705 affect virulence gene expression in Campylobacter jejuni. J Food Prot 76:1740–1746.
82. Chapman TA, Wu XY, Barchia I, Bettelheim KA, Driesen S, Trott D, Wilson M, Chin JJC. 2006. Comparison of virulence gene profiles of Escherichia coli strains isolated from healthy and diarrheic swine. Appl Environ Microbiol 72:4782–4795.
83. Fratamico PM, DebRoy C, Liu Y, Needleman DS, Baranzoni GM, Feng P. 2016. Advances in molecular serotyping and subtyping of Escherichia coli. Front Microbiol.
84. Kaper JB, Nataro JP, Mobley HLT. 2004. Pathogenic Escherichia coli. Nat Rev Microbiol.
85. Croxen MA, Finlay BB. 2010. Molecular mechanisms of Escherichia coli pathogenicity. Nat Rev Microbiol 8:26–38.
86. Donnenberg MS, Tacket CO, James SP, Losonsky G, Nataro JP, Wasserman SS, Kaper JB, Levine MM. 1993. Role of the eaeA gene in experimental enteropathogenic Escherichia coli infection. J Clin Invest 92:1412–1417.
87. Taneike I, Zhang H-M, Wakisaka-Saito N, Yamamoto T. 2002. Enterohemolysin operon of Shiga toxin-producing Escherichia coli: a virulence function of inlammatory cytokine production from human monocytes. FEBS Lett 524:219–224.
88. Beraldo LG, Borges CA, Maluta RP, Cardozo MV, Rigobelo EC, de Ávila FA. 2014. Detection

of Shiga toxigenic (STEC) and enteropathogenic (EPEC) Escherichia coli in dairy buffalo. Vet Microbiol 170:162–166.

89. Amin MA, Hashem HR, El-Mahallawy HS, Abdelrahman AA, Zaki HM, Azab MM. 2022. Characterization of enterohemorrhagic Escherichia coli from diarrhoeic patients with particular reference to production of Shiga-like toxin. Microb Pathog 166:105538.
90. Umpiérrez A, Ernst D, Fernández M, Oliver M, Casaux ML, Caffarena RD, Schild C, Giannitti F, Fraga M, Zunino P. 2021. Virulence genes of Escherichia coli in diarrheic and healthy calves. Rev Argent Microbiol 53:34–38.
91. McWilliams BD, Torres AG. 2014. Enterohemorrhagic Escherichia coli Adhesins. Microbiol Spectr 2.
92. Beltrametti F, Kresse AU, Guzma´n CA, Guzma´n G. 1999. Transcriptional Regulation of the esp Genes of Enterohemorrhagic Escherichia coliJOURNAL OF BACTERIOLOGY.
93. Navarro-Garcia F, Serapio-Palacios A, Ugalde-Silva P, Tapia-Pastrana G, Chaves-Duneas L. 2013. Actin Cytoskeleton Manipulation by Effector Proteins Secreted by Diarrheagenic Escherichia coli Pathotypes. Biomed Res Int 2013.
94. Jelčić I, Hüfner E, Schmidt H, Hertel C. 2008. Repression of the locus of the enterocyte effacement-encoded regulator of gene transcription of Escherichia coli O157:H7 by Lactobacillus reuteri culture supernatants is LuxS and strain dependent. Appl Environ Microbiol 74:3310–3314.
95. Sharp FC, Sperandio V. 2007. QseA directly activates transcription of LEE1 in enterohemorrhagic Escherichia coli. Infect Immun 75:2432–2440.
96. Beutin L, Strauch E, Zimmermann S, Kaulfuss S, Schaudinn C, Männel A, Gelderblom HR. 2005. Genetical and functional investigation of fliC genes encoding flagellar serotype H4 in wildtype strains of Escherichia coli and in a laboratory E. coli K-12 strain expressing flagellar antigen type H48. BMC Microbiol 5.
97. Zeinhom M, Tellez AM, Delcenserie V, El-Kholy AM, El-Shinawy SH, Griffiths MW. 2012. Yogurt containing bioactive molecules produced by Lactobacillus acidophilus la-5 exerts a protective effect against enterohemorrhagic Escherichia coli in mice. J Food Prot 75:1796–1805.
98. Medellin-Peña MJ, Wang H, Johnson R, Anand S, Griffiths MW. 2007. Probiotics affect virulence-related gene expression in Escherichia coli O157:H7. Appl Environ Microbiol 73:4259–4267.
99. Cadieux PA, Burton JP, Devillard E, Reid G. 2009. Lactobacillus by-products inhibit the growth and virulence of uropathogenic Escherichia coli. J Physiol Pharmacol 60:13–18.

100. Sherman PM, Johnson-Henry KC, Yeung HP, Ngo PSC, Goulet J, Tompkins TA. 2005. Probiotics reduce enterohemorrhagic Escherichia coli O157:H7- and enteropathogenic E. coli O127:H6-induced changes in polarized T84 epithelial cell monolayers by reducing bacterial adhesion and cytoskeletal rearrangements. Infect Immun 73:5183–5188.
101. Leccese Terraf MC, Juarez Tomás MS, Rault L, Le Loir Y, Even S, Nader-Macías MEF. 2017. In vitro effect of vaginal lactobacilli on the growth and adhesion abilities of uropathogenic Escherichia coli. Arch Microbiol 199:767–774.
102. Mack DR, Michail S, Wei S, McDougall L, Hollingsworth MA. 1999. Probiotics inhibit enteropathogenic E. coli adherence in vitro by inducing intestinal mucin gene expression. Am J Physiol - Gastrointest Liver Physiol 276.
103. Liévin-Le Moal V, Amsellem R, Servin AL, Coconnier MH. 2002. Lactobacillus acidophilus (strain LB) from the resident adult human gastrointestinal microflora exerts activity against brush border damage promoted by a diarrhoeagenic escherichia coli in human Enterocyte-like cells. Gut 50:803–811.
104. Atassi F, Brassart D, Grob P, Graf F, Servin AL. 2006. Vaginal Lactobacillus isolates inhibit uropathogenic Escherichia coli. FEMS Microbiol Lett 257:132–138.
105. Anand S, Mandal S, Singh KS, Patil P, Tomar SK. 2018. Synbiotic combination of Lactobacillus rhamnosus NCDC 298 and short chain fructooligosaccharides prevents enterotoxigenic Escherichia coli infection. Lwt 98:329–334.
106. Hirano J, Yoshida T, Sugiyama T, Koide N, Mori I, Yokochi T. 2003. The effect of Lactobacillus rhamnosus on enterohemorrhagic Escherichia coli infection of human intestinal cells in vitro. Microbiol Immunol 47:405–409.
107. Kim Y, Oh S, Park S, Seo JB, Kim SH. 2008. Lactobacillus acidophilus reduces expression of enterohemorrhagic Escherichia coli O157:H7 virulence factors by inhibiting autoinducer-2-like activity. Food Control 19:1042–1050.
108. Park H, Yeo S, Ji Y, Lee J, Yang J, Park S, Shin H, Holzapfel W. 2014. Autoinducer-2 associated inhibition by Lactobacillus sakei NR28 reduces virulence of enterohaemorrhagic Escherichia coli O157: H7. Food Control 45:62–69.
109. Chen YP, Lee TY, Hong WS, Hsieh HH, Chen MJ. 2013. Effects of Lactobacillus kefiranofaciens M1 isolated from kefir grains on enterohemorrhagic Escherichia coli infection using mouse and intestinal cell models. J Dairy Sci 96:7467–7477.
110. Chu H, Kang S, Ha S, Cho K, Park SM, Han KH, Sang KK, Lee HG, Seung HH, Yun CH, Choi Y. 2005. Lactobacillus acidophilus expressing recombinant K99 adhesive fimbriae has an inhibitory effect on adhesion of enterotoxigenic Escherichia coli. Microbiol Immunol

49:941–948.

111. Mangell P, Nejdfors P, Jeppsson B. 2002. Lactobacillus plantarum 299v Inhibits Intestinal Permeability. Dig Dis 47:511–516.
112. Asahara T, Nomoto K, Watanuki M, Yokokura T. 2001. Antimicrobial activity of intraurethrally administered probiotic Lactobacillus casei in a murine model of Escherichia coli urinary tract infection. Antimicrob Agents Chemother 45:1751–1760.
113. Yang Y, Galle S, Le MHA, Zijlstra RT, Gänzle MG. 2015. Feed fermentation with reuteran- and levan-producing Lactobacillus reuteri reduces colonization of weanling pigs by enterotoxigenic Escherichia coli. Appl Environ Microbiol 81:5743–5752.
114. Peng M, Reichmann G, Biswas D. 2015. Lactobacillus casei and its byproducts alter the virulence factors of foodborne bacterial pathogens. J Funct Foods 15:418–428.
115. Coconnier M-élène, Bernet M-F, Kernéis S, Chauvière G, Foirniat J, Servin AL. 1993. Inhibition of adhesion of enteroinvasive pathogens to human intestinal Caco-2 cells by Lactobacillus acidophilus strain LB decreases bacterial invasion. FEMS Microbiol Lett 110:299–306.
116. Bernet MF, Brassart D, Neeser JR, Servin AL. 1994. Lactobacillus acidophilus LA 1 binds to cultured human intestinal cell lines and inhibits cell attachment and cell invasion by enterovirulent bacteria. Gut 35:483–489.
117. Fonseca HC, de Sousa Melo D, Ramos CL, Dias DR, Schwan RF. 2021. Probiotic Properties of Lactobacilli and Their Ability to Inhibit the Adhesion of Enteropathogenic Bacteria to Caco-2 and HT-29 Cells. Probiotics Antimicrob Proteins 13:102–112.
118. Binyamin D, Nitzan O, Azrad M, Hamo Z, Koren O, Peretz A. 2021. The microbial diversity following antibiotic treatment of Clostridioides difficile infection. BMC Gastroenterol 21.
119. Lyerly DM, Krivan HC, Wilkins TD. 1988. Clostridium difficile: its disease and toxins. Clin Microbiol Rev 1:1–18.
120. Awad MM, Johanesen PA, Carter GP, Rose E, Lyras D. 2015. Clostridium difficile virulence factors: Insights into an anaerobic spore-forming pathogen. Gut Microbes 5:579–593.
121. Mani N, Lyras D, Barroso L, Howarth P, Wilkins T, Rood JI, Sonenshein AL, Dupuy B. 2002. Environmental response and autoregulation of Clostridium difficile TxeR, a sigma factor for toxin gene expression. J Bacteriol 184:5971–5978.
122. Trejo FM, Pérez PF, De Antoni GL. 2010. Co-culture with potentially probiotic microorganisms antagonises virulence factors of Clostridium difficile in vitro. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol 98:19–29.
123. Carasi P, Trejo FM, Pérez PF, De Antoni GL, Serradell M de los A. 2012. Surface proteins

from Lactobacillus kefir antagonize invitro cytotoxic effect of Clostridium difficile toxins. Anaerobe 18:135–142.

124. Najarian A, Sharif S, Griffiths MW. 2019. Evaluation of protective effect of Lactobacillus acidophilus La-5 on toxicity and colonization of Clostridium difficile in human epithelial cells in vitro. Anaerobe 55:142–151.
125. Yun B, Oh S, Griffiths MW. 2014. Lactobacillus acidophilus modulates the virulence of Clostridium difficile. J Dairy Sci 97:4745–4758.
126. Sagheddu V, Uggeri F, Belogi L, Remollino L, Brun P, Bernabè G, Moretti G, Porzionato A, Morelli L, Castagliuolo I, Elli M. 2020. The Biotherapeutic Potential of Lactobacillus reuteri Characterized Using a Target-Specific Selection Process. Front Microbiol 11.
127. Andersen KK, Strokappe NM, Hultberg A, Truusalu K, Smidt I, Mikelsaar RH, Mikelsaar M, Verrips T, Hammarström L, Marcotte H. 2016. Neutralization of Clostridium difficile toxin B mediated by engineered lactobacilli that produce single-domain antibodies. Infect Immun 84:395–406.
128. Gao X, Ma Y, Wang Z, Bai J, Jia S, Feng B, Jiang Y, Cui W, Tang L, Li Y, Wang L, Xu Y. 2019. Oral immunization of mice with a probiotic Lactobacillus casei constitutively expressing the α-toxoid induces protective immunity against Clostridium perfringens α-toxin. Virulence 10:166–179.
129. Parlet CP, Brown MM, Horswill AR. 2019. Commensal Staphylococci Influence Staphylococcus aureus Skin Colonization and Disease. Trends Microbiol. Elsevier Ltd.
130. Oogai Y, Matsuo M, Hashimoto M, Kato F, Sugai M, Komatsuzawa H. 2011. Expression of virulence factors by Staphylococcus aureus grown in Serum. Appl Environ Microbiol 77:8097–8105.
131. Giraudo AT, Calzolari A, Cataldi AA, Bogni C, Nagel R. 1999. The sae locus of Staphylococcus aureus encodes a two-component regulatory system. FEMS Microbiol Lett 177:15–22.
132. Argudín MÁ, Mendoza MC, Rodicio MR. 2010. Food Poisoning and Staphylococcus aureus Enterotoxins. Toxins (Basel) 2:1751–1773.
133. Zheng Y, Qin C, Zhang X, Zhu Y, Li A, Wang M, Tang Y, Kreiswirth BN, Chen L, Zhang H, Du H. 2020. The tst gene associated Staphylococcus aureus pathogenicity island facilitates its pathogenesis by promoting the secretion of inflammatory cytokines and inducing immune suppression. Microb Pathog 138.
134. Tang A, Caballero AR, Bierdeman MA, Marquart ME, Foster TJ, Monk IR, O'callaghan RJ. 2019. Staphylococcus aureus Superantigen-Like Protein SSL1: A Toxic Protease. Pathogens

8.

135. Atkins KL, Burman JD, Chamberlain ES, Cooper JE, Poutrel B, Bagby S, Jenkins ATA, Feil EJ, van den Elsen JMH. 2008. S. aureus IgG-binding proteins SpA and Sbi: Host specificity and mechanisms of immune complex formation. Mol Immunol 45:1600–1611.
136. Lakhundi S, Zhang K. 2018. Methicillin-Resistant Staphylococcus aureus: Molecular Characterization, Evolution, and Epidemiology. Clin Microbiol Rev 31.
137. Omidi M, Firoozeh F, Saffari M, Sedaghat H, Zibaei M, Khaledi A. 2020. Ability of biofilm production and molecular analysis of spa and ica genes among clinical isolates of methicillin-resistant Staphylococcus aureus. BMC Res Notes 13–19.
138. Alp F, Ssf S, Crg S, Mvp L. 2019. Lactobacillus rhamnosus versus Staphylococcus aureus: influence on growth and expression of virulence factors. J Dent Maxillofac Res 2:29–33.
139. Jabbar H, Hala F, Radeef M. 2011. Capability of Lactobacillus acidophilus supernatant to inhibit production of lipase from methicillin-resistant Staphylococcus aureus. J Univ Anbar Pure Sci 5.
140. Melo TA, Dos Santos TF, De Almeida ME, Junior LAGF, Andrade EF, Rezende RP, Marques LM, Romano CC. 2016. Inhibition of Staphylococcus aureus biofilm by Lactobacillus isolated from fine cocoa. BMC Microbiol 16:1–9.
141. Ramezani M, Zainodini N, Hakimi H, Zarandi ER, Bagheri V, Bahramabadi R, Zare-Bidaki M. 2019. Cell-free culture supernatants of lactobacilli modify the expression of virulence factors genes in staphylococcus aureus. Jundishapur J Microbiol 12.
142. Laughton JM, Devillard E, Heinrichs DE, Reid G, Mccormick JK. 2006. Inhibition of expression of a staphylococcal superantigen-like protein by a soluble factor from Lactobacillus reuteri. Microbiology 152:1155–1157.
143. Reid G, Tieszer C. 1994. Use of lactobacilli to reduce the adhesion of Staphylococcus aureus to catheters. Int Biodeterior Biodegrad 34:73–83.
144. Gan BS, Kim J, Reid G, Cadieux P, Howard JC. 2002. Lactobacillus fermentum RC-14 Inhibits Staphylococcus aureus Infection of Surgical Implants in Rats . J Infect Dis 185:1369–1372.
145. Younes JA, van der Mei H, ven den Heuvel E, Busscher HJ, Reid G. 2012. Adhesion Forces and Coaggregation between Vaginal Staphylococci and Lactobacilli. PLoS One 7.
146. Bouchard DS, Rault L, Berkova N, Le Loir Y, Even S. 2013. Inhibition of Staphylococcus aureus invasion into bovine mammary epithelial cells by contact with live Lactobacillus casei. Appl Environ Microbiol 79:877–885.
147. Ren D, Li C, Qin Y, Yin R, Li X, Tian M, Du S, Guo H, Liu C, Zhu N, Sun D, Li Y, Jin N. 2012. Inhibition of Staphylococcus aureus adherence to Caco-2 cells by lactobacilli and cell

surface properties that influence attachment. Anaerobe 18:508–515.

148. Chang WL, Yeh YC, Sheu BS. 2018. The impacts of H. pylori virulence factors on the development of gastroduodenal diseases. J Biomed Sci 25:1–9.

149. Yan J, Liang S-H, Mao Y-F, Li li-W, Li S-P. 2003. Construction of expression system for flaA and flaB genes of Helicobacter pylori and determination of immunoreactivity and antigenicity of recombinant proteins. World J Gastroenterol 9:2240–2250.

150. Liévin Le Moal V, Fayol-Messaoudi D, Servin AL. 2013. Compound(s) secreted by Lactobacillus casei strain Shirota YIT9029 irreversibly and reversibly impair the swimming motility of Helicobacter pylori and Salmonella enterica serovar Typhimurium, respectively. Microbiol (United Kingdom) 159:1956–1971.

151. Urrutia-Baca VH, Escamilla-García E, de la Garza-Ramos MA, Tamez-Guerra P, Gomez-Flores R, Urbina-Ríos CS. 2018. In Vitro Antimicrobial Activity and Downregulation of Virulence Gene Expression on Helicobacter pylori by Reuterin. Probiotics Antimicrob Proteins 10:168–175.

152. Ki M-R, Ghim S-Y, Hong I-H, Park J-K, Hong K-S, Ji A-R, Jeong K-S. 2010. In Vitro Inhibition of Helicobacter pylori Growth and of Adherence of cagA-Positive Strains to Gastric Epithelial Cells by Lactobacillus paraplantarum KNUC25 Isolated from Kimchi. J Med Food 13:629–634.

153. Ryan KA, O'Hara AM, Van Pijkeren JP, Douillard FP, O'Toole PW. 2009. Lactobacillus salivarius modulates cytokine induction and virulence factor gene expression in Helicobacter pylori. J Med Microbiol 58:996–1005.

154. Pena JA, Rogers AB, Ge Z, Ng V, Li SY, Fox JG, Versalovic J. 2005. Probiotic Lactobacillus spp. Diminish Helicobacter hepaticus-InducedInflammatory Bowel Disease in Interleukin-10-Deficient Mice. Infect Immun 73:912–920.

155. Gotteland M, Poliak L, Cruchet S, Brunser O. 2005. Effect of regular ingestion of Saccharomyces boulardii plus inulin or Lactobacillus acidophilus LB in children colonized by Helicobacter pylori. Acta Paediatr Int J Paediatr 94:1747–1751.

156. Jamal M, Ahmad W, Andleeb S, Jalil F, Imran M, Nawaz MA, Hussain T, Ali M, Rafiq M, Kamil MA. 2018. Bacterial biofilm and associated infections. J Chinese Med Assoc. Elsevier Ltd.

157. Veesenmeyer JL, Hauser AR, Lisboa T, Rello J. 2009. Pseudomonas aeruginosa virulence and therapy: Evolving translational strategies. Crit Care Med 37:1777–1786.

158. Sultan M, Arya R, Kim KK. 2021. Roles of two-component systems in pseudomonas aeruginosa virulence. Int J Mol Sci.

159. Alexandre Y, Le Berre R, Barbier G, Le Blay G. 2014. Screening of Lactobacillus spp. for the prevention of Pseudomonas aeruginosa pulmonary infections. BMC Microbiol 14:1–10.
160. Berríos P, Fuentes JA, Salas D, Carreño A, Aldea P, Fernández F, Trombert AN. 2018. Inhibitory effect of biofilm-forming Lactobacillus kunkeei strains against virulent Pseudomonas aeruginosa in vitro and in honeycomb moth (Galleria mellonella) infection model. Benef Microbes 9:257–268.
161. Valdéz JC, Peral MC, Rachid M, Santana M, Perdigón G. 2005. Interference of Lactobacillus plantarum with Pseudomonas aeruginosa in vitro and in infected burns: The potential use of probiotics in wound treatment. Clin Microbiol Infect 11:472–479.
162. Cui T, Bai F, Sun M, Lv X, Li X, Zhang D, Du H. 2020. Lactobacillus crustorum ZHG 2-1 as novel quorum-quenching bacteria reducing virulence factors and biofilms formation of Pseudomonas aeruginosa. Lwt 117:108696.
163. Wu MC, Lin TL, Hsieh PF, Yang HC, Wang JT. 2011. Isolation of genes involved in biofilm formation of a klebsiella pneumoniae strain causing pyogenic liver abscess. PLoS One 6.
164. Maldonado NC, Silva De Ruiz C, Cecilia M, Nader-Macias ME. 2007. A simple technique to detect Klebsiella biofilm-forming-strains. Inhibitory potential of Lactobacillus fermentum CRL 1058 whole cells and products. Commun Curr Res Educ Top Trends Appl Microbiol 52–59.
165. Al-Mathkhury HJF, Aded Assa SD. 2012. Inhibitory Effect of Lactobacilli Filtrate on Klebsiella Pneumoniae Biofilm. Iraqui Postgrad Med J 11:168–179.
166. Lemos JA, Palmer SR, Zeng L, Wen ZT, Kajfasz JK, Freires IA, Abranches J, Brady LJ. 2018. The Biology of Streptococcus mutans. Microbiol Spectr 7.
167. Krzyściak W, Jurczak A, Kościelniak D, Bystrowska B, Skalniak A. 2014. The virulence of Streptococcus mutans and the ability to form biofilms. Eur J Clin Microbiol Infect Dis 33:499–515.
168. Salehi R, Salehi A, Savabi O, Tahmourespour A, Eslami G, kamali S, Kazemi M. 2014. Effects of Lactobacillus reuteri-derived biosurfactant on the gene expression profile of essential adhesion genes (gtfB, gtfC and ftf) of Streptococcus mutans. Adv Biomed Res 3:169.
169. Tahmourespour A, Salehi R, Kasra Kermanshahi R. 2011. Lactobacillus acidophilus-derived biosurfactant effect on GTFB and GTFC expression level in Streptococcus mutans biofilm cells. Brazilian J Microbiol 42:330–339.
170. Tahmourespour A, Salehi R, Kermanshahi RK, Eslami G. 2011. The anti-biofouling effect of Lactobacillus fermentum-derived biosurfactant against Streptococcus mutans. Biofouling 27:385–392.

171. Ahmed A, Dachang W, Lei Z, Jianjun L, Juanjuan Q, Yi X. 2014. Effect of Lactobacillus species on Streptococcus mutans biofilm formation. Pak J Pharm Sci 27:1523–1528.
172. Wen ZT, Yates D, Ahn SJ, Burne RA. 2010. Biofilm formation and virulence expression by Streptococcus mutans are altered when grown in dual-species model. BMC Microbiol 10.
173. Wasfi R, Abd El-Rahman OA, Zafer MM, Ashour HM. 2018. Probiotic Lactobacillus sp. inhibit growth, biofilm formation and gene expression of caries-inducing Streptococcus mutans. J Cell Mol Med 22:1972–1983.
174. Söderling EM, Marttinen AM, Haukioja AL. 2011. Probiotic Lactobacilli interfere with Streptococcus mutans biofilm formation in Vitro. Curr Microbiol 62:618–622.
175. Wu CC, Lin CT, Wu CY, Peng WS, Lee MJ, Tsai YC. 2015. Inhibitory effect of Lactobacillus salivarius on Streptococcus mutans biofilm formation. Mol Oral Microbiol 30:16–26.
176. Fiedler T, Koller T, Kreikemeyer B. 2015. Streptococcus pyogenes biofilms - formation, biology, and clinical relevance. Front Cell Infect Microbiol 5.
177. Molloy EM, Cotter PD, Hill C, Mitchell DA, Ross RP. 2011. Streptolysin S-like virulence factors: the continuing sagA. Nat Rev Microbiol 9:670–681.
178. Saroj SD, Maudsdotter L, Tavares R, Jonsson AB. 2016. Lactobacilli interfere with streptococcus pyogenes hemolytic activity and adherence to host epithelial cells. Front Microbiol 7:1–8.
179. Rizzo A, Losacco A, Carratelli CR, Di Domenico M, Bevilacqua N. 2013. Lactobacillus plantarum reduces Streptococcus pyogenes virulence by modulating the IL-17, IL-23 and Toll-like receptor 2/4 expressions in human epithelial cells. Int Immunopharmacol 17:453–461.
180. Spurbeck RR, Arvidson CG. 2011. Lactobacilli at the front line of defense against vaginally acquired infections. Future Microbiol.
181. Redondo-Lopez V, Cook RL, Sobel JD. 1990. Emerging Role of Lactobacilli in the Control and Maintenance of the Vaginal Bacterial MicrofloraREVIEWS OF INFECTIOUS DISEASES.
182. Reid G, Cook RL, Bruce AW. 1987. Examination of strains of lactobacilli for properties that may influence bacterial interference in the urinary tract. J Urol 138:330–335.
183. Boris S, Suarez juan E, Vasques F, Barbés C. 1998. Adherence of Human Vaginal Lactobacilli to Vaginal Epithelial Cells and Interaction with Uropathogens. Infect Immun 66:1985–1989.
184. Kmet V, Lucchino F. 1997. Aggregationpromoting factor in human vaginal Lactobacillus strains. FEMS Immunol Med Microbiol 19:111–114.
185. Hawthorn LA, Reid G. 1990. Exclusion of uropathogen adhesion to polymer surfaces by Lactobacillus acidophilus.

186. Mielczarek E, Blaszkowska J. 2016. Trichomonas vaginalis: pathogenicity and potential role in human reproductive failure. Infection 44:447–458.
187. Quillin SJ, Seifert HS. 2018. Neisseria gonorrhoeae host adaptation and pathogenesis. Nat Rev Microbiol.
188. Castro J, Martins AP, Rodrigues ME, Cerca N, Onderdonk AB. 2018. Lactobacillus crispatus represses vaginolysin expression by BV associated Gardnerella vaginalis and reduces cell cytotoxicity. Anaerobe 50:60–63.
189. Phukan N, Parsamand T, Brooks AES, Nguyen TNM, Simoes-Barbosa A. 2013. The adherence of Trichomonas vaginalis to host ectocervical cells is influenced by lactobacilli. Sex Transm Infect 89:455–459.
190. Spurbeck RR, Arvidson CG. 2010. Lactobacillus jensenii surface-associated proteins inhibit Neisseria gonorrhoeae adherence to epithelial cells. Infect Immun 78:3103–3111.
191. Chang TL-Y, Chang C-H, Simpson DA, Xu Q, Martin PK, Lagenaur LA, Schoolnik GK, Ho DD, Hillier SL, Holodniy M, Lewicki JA, Lee PP. 2003. Inhibition of HIV infectivity by a natural human isolate of Lactobacillus jensenii engineered to express functional two-domain CD4.
192. Atassi F, Brassart D, Grob P, Graf F, Servin AL. 2006. Lactobacillus strains isolated from the vaginal microbiota of healthy women inhibit Prevotella bivia and Gardnerella vaginalis in coculture and cell culture. FEMS Immunol Med Microbiol 48:424–432.
193. Atassi F, Pho Viet Ahn DL, Lievin-Le Moal V. 2019. Diverse Expression of Antimicrobial Activities Against Bacterial Vaginosis and Urinary Tract Infection Pathogens by Cervicovaginal Microbiota Strains of Lactobacillus gasseri and Lactobacillus crispatus. Front Microbiol 10:1–13.
194. Atassi F, Brassart D, Grob P, Graf F, Servin AL. 2006. In vitro antibacterial activity of Lactobacillus helveticus strain KS300 against diarrhoeagenic, uropathogenic and vaginosis-associated bacteria. J Appl Microbiol 101:647–654.
195. Maras B, Maggiore A, Mignogna G, D'Erme M, Angiolella L. 2021. Hyperexpression of CDRs and HWP1 genes negatively impacts on Candida albicans virulence. PLoS One 16.
196. Mayer FL, Wilson C, Hube B. 2013. Candida albicans pathogenicity mechanisms. Virulence 4:119–128.
197. De Seta F, Parazzini F, De Leo R, Banco R, Maso GP, De Santo D, Sartore A, Stabile G, Inglese S, Tonon M, Restaino S. 2014. Lactobacillus plantarum P17630 for preventing Candida vaginitis recurrence: A retrospective comparative study. Eur J Obstet Gynecol Reprod Biol 182:136–139.

198. da Silva Dantas A, Lee KK, Raziunaite I, Schaefer K, Wagener J, Yadav B, Gow NA. 2016. Cell biology of Candida albicans–host interactions. Curr Opin Microbiol. Elsevier Ltd.
199. Grin PM, Kowalewska PM, Alhazzani W, Fox-robichaud AE. 2013. Lactobacillus for preventing recurrent urinary tract infections in women: meta-analysis. Can J Urol 20:6607–6614.
200. Itapary dos Santos C, Ramos Franca Y, Duarte Lima Campos C, Quaresma Bomfim MR, Oliveira Melo B, Assuncao Holanda R, Santos VL, Gomes Monteiro S, Buozzi Moffa E, Souza Monteiro A, Andrade Monteiro C, Monteiro-Nero V. 2019. Antifungal and Antivirulence Activity of Vaginal Lactobacillus Spp. Products against Candida Vaginal Isolates. Pathogens 8.
201. Santos CMA, Pires MCV, Leão TL, Hernández ZP, Rodriguez ML, Martins AKS, Miranda LS, Martins FS, Nicoli JR. 2016. Selection of Lactobacillus strains as potential probiotics for vaginitis treatment. Microbiol (United Kingdom) 162:1195–1207.
202. Matsuda Y, Cho O, Sugita T, Ogishima D, Takeda S. 2018. Culture Supernatants of Lactobacillus gasseri and L. crispatus Inhibit Candida albicans Biofilm Formation and Adhesion to HeLa Cells. Mycopathologia 183:691–700.
203. Graf K, Last A, Gratz R, Allert S, Linde S, Westermann M, Gröger M, Mosig AS, Gresnigt MS, Hube B. 2019. Keeping Candida commensal: How lactobacilli antagonize pathogenicity of Candida albicans in an in vitro gut model. DMM Dis Model Mech 12.
204. Allonsius CN, Vandenheuvel D, Oerlemans EFM, Petrova MI, Donders GGG, Cos P, Delputte P, Lebeer S. 2019. Inhibition of Candida albicans morphogenesis by chitinase from Lactobacillus rhamnosus GG. Sci Rep 9:1–12.
205. Oliveira VMC, Santos SSF, Silva CRG, Jorge AOC, Leão MVP. 2016. Lactobacillus is able to alter the virulence and the sensitivity profile of Candida albicans. J Appl Microbiol 121:1737–1744.
206. Mailänder-Sánchez D, Braunsdorf C, Grumaz C, Müller C, Lorenz S, Stevens P, Wagener J, Hebecker B, Hube B, Bracher F, Sohn K, Schaller M. 2017. Antifungal defense of probiotic Lactobacillus rhamnosus GG is mediated by blocking adhesion and nutrient depletion. PLoS One 12:1–19.
207. Martinez RCR, Seney SL, Summers KL, Nomizo A, De Martinis ECP, Reid G. 2009. Effect of Lactobacillus rhamnosus GR-1 and Lactobacillus reuteri RC-14 on the ability of Candida albicans to infect cells and induce inflammation. Microbiol Immunol 53:487–495.
208. Niu XX, Li T, Zhang X, Wang SX, Liu ZH. 2017. Lactobacillus crispatus modulates vaginal epithelial cell innate response to Candida albicans. Chin Med J (Engl) 130:273–279.

209. Rizzo A, Losacco A, Carratelli CR. 2013. Lactobacillus crispatus modulates epithelial cell defense against Candida albicans through Toll-like receptors 2 and 4, interleukin 8 and human β-defensins 2 and 3. Immunol Lett 156:102–109.
210. Parolin C, Marangoni A, Laghi L, Foschi C, Palomino RAÑ, Calonghi N, Cevenini R, Vitali B. 2015. Isolation of vaginal lactobacilli and characterization of anti-candida activity. PLoS One 10:1–17.
211. Ahmad A. 2013. Probiotic Lactobacillus affects the expression of virulence in Candida albicans. J Microb Biochem Technol 5:5948.
212. Wang S, Wang Q, Yang E, Yan L, Li T, Zhuang H. 2017. Antimicrobial compounds produced by vaginal Lactobacillus crispatus are able to strongly inhibit Candida albicans growth, hyphal formation and regulate virulence-related gene expressions. Front Microbiol 8:1–11.
213. Ribeiro FC, de Barros PP, Rossoni RD, Junqueira JC, Jorge AOC. 2017. Lactobacillus rhamnosus inhibits Candida albicans virulence factors in vitro and modulates immune system in Galleria mellonella. J Appl Microbiol 122:201–211.
214. Rossoni RD, Fuchs BB, De Barros PP, Dos Santos Velloso M, Jorge AOC, Junqueira JC, Mylonakis E. 2017. Lactobacillus paracasei modulates the immune system of Galleria mellonella and protects against Candida albicans infection. PLoS One 12:1–17.
215. Vilela SFG, Barbosa JO, Rossoni RD, Santos JD, Prata MCA, Anbinder AL, Jorge AOC, Junqueira JC. 2015. Lactobacillus acidophilus ATCC 4356 inhibits biofilm formation by C. Albicans and attenuates the experimental candidiasis in Galleria mellonella. Virulence 6:29–39.
216. De Montijo-Prieto S, Moreno E, Bergillos-Meca T, Lasserrot A, Ruiz-López MD, Ruiz-Bravo A, Jiménez-Valera M. 2015. A Lactobacillus plantarum strain isolated from kefir protects against intestinal infection with Yersinia enterocolitica O9 and modulates immunity in mice. Res Microbiol 166:626–632.
217. Felipe EMM, Fanny MJ, Isabela A, Celia RG, Mariella VPL, Silvana SFDS. 2016. Relationship between the probiotic Lactobacillus rhamnosus and Enterococcus faecalis during the biofilm formation. African J Microbiol Res 10:1182–1186.
218. Reda FM. 2014. Detoxification of enterotoxigenic Bacillus cereus (JX455159) isolated from meat by a local strain of Lactobacillus plantarum (JX282192). Ann Microbiol 64:287–296.
219. Kasra-Kermanshahi R, Vahedi-Shahandashti R, Shokouhfard M, Ghadam P, Feizabadi MM, Teimourian S. 2017. Antagonistic activities of some probiotic lactobacilli culture supernatant on Serratia marcescens swarming motility and antibiotic resistance.
220. Nissen L, Sgorbati B, Biavati B, Belibasakis GN. 2014. Lactobacillus salivarius and L. gasseri

down-regulate Aggregatibacter actinomycetemcomitans exotoxins expression. Ann Micriobiol 64:611–617.

221. EFSA. 2007. Introduction of a Qualified Presumption of Safety (QPS) approach for assessment of selected microorganisms referred to EFSA - Opinion of the Scientific Committee. EFSA J 5:1–16.

222. Colautti A, Arnoldi M, Comi G, Iacumin L. 2021. Antibiotic resistance and virulence factors in lactobacilli: Something to carefully consider. Food Microbiol 103:103934.

1.6 Supplementary material

Supplementary table 1: Lactobacilli strains inhibiting *Listeria monocytogenes* virulence factors

Lactobacilli	Strain	Effect on	Reference
L. salivarius	UCC118	Inflammation response	(58)
L. acidophilus	NCDO 1205	Inflammation response	(58)
	ACCC11073	Cytokines level, translocation to organs, and LLO, InlA, InlB, Ami, and flagellin production	(59)
	LA 1	Adhesion and invasion	(60)
	LB		(61)
L. plantarum	LB95	Invasiveness	(62)
	CICC 6257	*sigB, hly, inlA, inlB,* and *prfA* expression	(63)
	B-4496	Adhesion, invasion and virulence gene expression	(64)
	CICC21863	Cytokines level, translocation to organs, and LLO, InlA, InlB, Ami, and flagellin production	(59)
L. paracasei	Recombinant LAP expressing	Adhesion and invasion	(65)
	CNCM I-3689	Infection	(66)
L casei	Recombinant InlA InlB expressing	Adhesion, invasion and citotoxicity	(67)
	BL23	Infection	(66)
	CFCS1 CFCS2	*fbp* and *iap* expression	(1)
L. rhamnosus	GG	Adhesion and invasion	(68)
L. sakei	2A	Citotoxicity	(69)
	1	Adhesion	(70)
			(71)
		Hemolytic activity	(72)
			(73)
L. fermentum	B-1840	Adhesion, invasion and virulence gene expression	(64)
L. reuteri	B-14172	Adhesion, invasion and virulence gene expression	(64)
L. brevis	MF179529	Translocation to organs	(74)

Supplementary table 2: Lactobacilli inhibiting *Salmonella* spp. virulence factors

Lactobacilli	Strain	*Salmonella* spp.	Effect on	Reference
L. bulgaricus	NRRL B548	*S.* Enteritidis *S.* Heidelberg *S.* Typhimurium	*sipA, sipB, sopB, spvB, hilA, hilD*, and *invH* expression	(12)
		S. Typhimurium	*hilA, hilD*, *hilC,* and *sipC* expression	(75)
L. casei	-	*S.* Typhimurium	*hilA, hilD*, *hilC,* and *sipC* expression	(75)
	Shirota		Adhesion and invasion	(76)
	Shirota YIT9029		Swimming motility	(77)
	Recombinant LC-CLA		Biofilm formation and interaction with the host	(78)
	Recombinant LC-CLA		Physicochemical properties, interaction with the host, *invG, invH, prgK, hilA, hilC, hilD*, and *invF* expression	(79)
	ATCC 334			
	-		Invasion and translocation to organs	(80)
	CFCS1 CFCS2		*nmpC* expression	(1)
	-	*S.* Javiana	Citotoxicity and invasiveness	(81)
L. amylovorus	CL12	*S.* Typhimurium	*hilA, hilC, hilD, sopB, sopD, sopE2, sipA, avrA, sptP* expression	(82) (83)
	DCE 471			(76)
L. rhamnosus	-	*S.* Typhimurium	*hilA, hilD*, *hilC,* and *sipC* expression	(75)
	L2		*hilA, hilC, hilD, sopB, sopD, sopE2, sipA, avrA, sptP* expression	(82)
	L3			(82)
	LB2			(83)
	LB4			
	GG		Invasiveness	(84)
			Adhesion and invasion	(76)
			Growth	(85)
	-	*S.* Javiana	Citotoxicity and invasiveness	(81)
	NRRLB442	*S.* Enteritidis *S.* Heidelberg	*sipA, sipB, sopB, spvB, hilA, hilD,* and *invH* expression	(12)
L. acidophilus	-	*S.* Typhimurium	*hilA, hilD*, *hilC,* and *sipC* expression	(75)
	CL10		*hilA, hilC, hilD, sopB, sopD, sopE2, sipA, avrA, sptP* expression	(82)
	CL10			(83)
	-		*invA, avrA, hilA, ssrB,* and *sopD* expression	(86)
	IBB 801		Adhesion and invasion	(76)
	LB		permeabilization of the membrane, sensitivity to sodium dodecyl sulfate and death	(87)
	LA 1		Adhesion and invasion	(61) (60)
	-	*S.* Javiana	Citotoxicity and invasiveness	(81)
L. brevis	CCMA 1284	*S. Enteritidis*		(88)

Unknown	-	*S.* Enteritidis	β-galactosidase activity and *hilA* expression	(89)
L. crispatus	ALB11	*S.* Typhimurium	*hilA, hilC, hilD, sopB, sopD, sopE2, sipA, avrA, sptP* expression	(82) (83)
L. johnsonii	La1	*S.* Typhimurium	Adhesion and invasion	(76)
L. plantarum	ZS2058	*S.* Typhimurium	*invA, avrA, hilA, ssrB,* and *sopD* expression	(90)
	S8 S66		*hilA, hilC, hilD, sopB, sopD, sopE2, sipA, avrA, sptP* expression	(83)
	C4 C7 C8 B2a B10 B11 L4 L36 L37 L38 L39		Resistance to antibiotics, adhesion and citotoxicity	(91)
	ACA-DC 287		Adhesion and invasion	(76)
	-		Adhesion	(92)
	S8		Pro-inflammatory cytokine response	(93)
	CCMA 0359 CCMA 0743	*S. Enteritidis*	Adhesion	(88)
L. salivarius	ALB2 ALB6	*S.* Typhimurium	*hilA, hilC, hilD, sopB, sopD, sopE2, sipA, avrA, sptP* expression	(83)
	ALB2 ALB6 ALB7 ALB10			(82)
	SG1		Adhesion	(92)
L. reuteri	CL9	*S.* Typhimurium	*hilA, hilC, hilD, sopB, sopD, sopE2, sipA, avrA, sptP* expression	(82)
	CL9 S64 K67			(83)
	S64		Pro-inflammatory cytokine response	(93)
L. zeae	LB1	*S.* Typhimurium	*hilA, hilC, hilD, sopB, sopD, sopE2, sipA, avrA, sptP* expression	(82) (83)
	LB2		Pro-inflammatory cytokine response	(93)
L. delbrueckii var delbrueckii	-	*S.* Typhimurium	Adhesion	(92)
L. paracasei	DUP-13076	*S.* Enteritidis *S.* Heidelberg	*sipA, sipB, sopB, spvB, hilA, hilD,* and *invH* expression	(12)
	IBB2588	*S.* Enteritidis	Adhesion	(94)

	CCMA 0504 CCMA 0505		(88)

Supplementary table 3: Lactobacilli inhibiting *Campylobacter jejuni* virulence factors

Lactobacilli	Strain	Effect on	Reference
L. salivarius	AH102	Internalization	(95)
	-	Growth, *flaA, flaB, flhA, ciaB, luxS* expression, phagocytosis	(17)
L. johnsonii	-	Growth, *flaA, flaB, flhA, ciaB, luxS* expression, phagocytosis	(17)
L. reuteri	-	Growth, phagocytosis	(17)
L. crispatus	-	Growth, *flaA, flaB, flhA, ciaB, luxS* expression, phagocytosis	(17)
L. gasseri	-	Growth, *flaA, flaB, flhA, ciaB, luxS* expression, phagocytosis	(17)
L. helveticus	R0052	Internalization	(95)
L. casei	*recombinant mcra* expressing	Adhesion and *cadF, cdtB, ciaB,* and *flaB* expression	(96)
L. acidophilus	La-5	*luxS* expression	(97)
L. rhamnosus	R0011	Internalization	(95)

Supplementary table 4: Lactobacilli strains inhibiting *Escherichia coli* virulence factors

Lactobacilli	Strain	*Escherichia* spp.	Effect on	Reference
L. reuteri	ATCC 55730	EHEC	*ler* expression	(21)
	RC-14	UPEC	Adhesion and virulence gene expression	(98)
	CRL 1324	UPEC	Adhesion and internalization	(99)
	TMW1.656	ETEC	Toxins production	(100)
	LTH5794		Toxins production	
L. plantarum	299v	-	Internalization	(101)
		EPEC E2348/69	Adhesion	(102)
		EHEC CL8	Adhesion	
	CCMA 0359	EPEC CDC	Adhesion	(88)
	CCMA 0743	055	Adhesion	
L. acidophilus	La-5	EHEC	Colonization and TNF-α production	(22)
		EHEC O157	*tir*, *espA*, *fliC*, *espD*, *luxS*, *eaeA*, *ler, hylB,* and *qseA* expression	(23)
	R0052	EHEC O157	Adhesion	(103)
		EPEC E2348/69	Adhesion	
	A4	EHEC	Shiga-like Toxin 2 activity	(104)
	K99	ETEC	Adhesion	(105)
	LA 1	EPEC	Adhesion and invasion	(60)
	LB	EPEC	Adhesion and invasion	(61)
		DAEC	Expression of virulence genes	(106)
L. sakei	NR28	EHEC	Biofilm formation, AI-2 expression and adhesion	(104)
	NR28		AI-2 production	(107)
L. casei	Recombinant LC-CLA	EHEC	Adhesion and invasion	(78)
	CFCS1		*eaeA* expression	(1)
	CFCS2		*eaeA* expression	
	Shirota	-	Growth rate and inflammatory response	(108)
L. rhamnosus	R0011	EHEC O157 EPEC E2348/69	Adhesion	(103)
	GG	EPEC E2348/69		(102)
		EHEC CL8		
	NCDC 298	ETEC		(109)
	-		Internalization	(110)
L. kefiranofaciens	M1	EHEC	Immune response	(111)
L. paracasei	CCMA 0504 CCMA 0505	EPEC	Adhesion	(88)

L. gasseri *L. jensenii*	KS120.1 KS124.3 KS119.1 KS121.1	DAEC	Adhesion and internalization	(112)

Supplementary table 5: Lactobacilli strains inhibiting *Clostridium* spp. virulence factors

***Clostridium* spp.**	**Lactobacilli**	**Strain**	**Effect on**	**Reference**
C. difficile	*L. acidophilus*	ATCC 314	TcdA and TcdB production	(113)
	L. brevis	ATCC 8287	TcdA and TcdB production	(113)
	L. plantarum	CIDCA 83114	TcdA and TcdB production	(113)
	L. acidophilus	La-5	Adhesion	(114)
		GP1B	*luxS, tcdA, tcdB,* and *txeR* expression	(29)
	L. kefir	CIDCA 8348	TcdA and TcdB production	(113)
		CIDCA 8344	Citotoxicity	(115)
		CIDCA 83111		
		CIDCA 83113		
		CIDCA 83115		
		CIDCA 8321		
		CIDCA 8345		
		CIDCA 8348		
		JCM 5818		
		ATCC 8007		
	L. reuteri	LMG P-27481	Colonization and toxins production	(116)
	L. paracasei	Recombinant anti-TcdBVHH fragment-expressing	Citotoxicity	(117)
	L. casei	DSMZ 20011	TcdA and TcdB production	(113)
C. perfrigens		Recombinant pPG-α 393	Citokines and interferon γ production	(118)

Supplementary table 6: Lactobacilli strains inhibiting *Staphylococcus aureus* virulence factors

Lactobacilli	Strain	Effect on	Reference
L. acidophilus	-	Biofilm formation and antibiotic resistance	(119)
	76 T-13	Adhesion	(120)
	ATCC 4356	Adhesion	(121)
L. fermentum	TCUESC01	*icaA* and *icaR* expression	(122)
	ATCC 9338 B-54	*sea, sae, agrA, tst, spa,* and *spi* expression	(123)
	RC-14	Adhesion	(124)
L. plantarum	TCUESC02	Growth	(122)
	CGMCC 1.557	Adhesion	(121)
L. reuteri	ATCC 23272	*sea, sae, agrA, tst, spa,* and *spi* expression	(123)
	RC-14	SSL 1 production	(125)
L. casei	36	Adhesion	(120)
	ATCC 393 BL23 CIRM-BIA 1542	Internalization	(126)
L. crispatus	33820	Coaggregation	(127)
L. plantarum	ATCC 8014	*sea, sae, agrA, tst, spa,* and *spi* expression	(123)
L. jensenii	RC-28	Coaggregation	(127)
L. rhamnosus	GR-1	Adhesion	(120)
	ATCC 1465	Biofilm formation	(128)

Supplementary table 7: Lactobacilli strains inhibiting *Helicobacter* spp. virulence factors

***Helicobacter* spp.**	**Lactobacilli**	**Strain**	**Effect on**	**Reference**
H. pylori	*L. salivarius*	UCC118 UCC119	*Cag* expression and interleukin immune response	(129)
	L. acidophilus	LB	Viability	(130)
	L. paraplantarum	KNUC25	Adhesion	(131)
	L. casei Shirota	YIT9029	Swimming motility	(77)
	L. reuteri	ATCC 55730	*flaA* and *vacA* expression	(132)
H. hepaticus		6798	Interleukin and chemokin response	(133)
	L. paracasei	1602	Interleukin and chemokin response	(133)

Supplementary table 8: Lactobacilli strains inhibiting *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, *Streptococcus mutans*, and *Streptococcus pyogenes* virulence factors

Lactobacilli	Strain	Pathogen	Effect on	Reference
L. fermentum	-	*Pseudomonas aeruginosa*	Biofilm and elastase production	(134)
	CRL 1058	*Klebsiella pneumoniae*	Adhesion	(135)
	-		Replication inside biofilm	(136)
	ATCC 9338	*Streptococcus mutans*	*gtfB* and *gtfC* expression	(137)
L. zeae	-	*Pseudomonas aeruginosa*	Biofilm and elastase production	(134)
L. paracasei	-	*Pseudomonas aeruginosa*	Biofilm and elastase production	(134)
C. crustorum	ZHG 2-1	*Pseudomonas aeruginosa*	*lasI/R* and *rhlI/R* expression	(45)
L. salivarius	ATCC 11741	*Streptococcus mutans*	Biofilm formation	(138)
	K35 K43		*gtfB, gtfC, gtfD* expression	(139)
	LMG9477	*Streptococcus pyogenes*	Adhesion, hemolytic activity and *sag* expression	(140)
L. rhamnosus	GG ATCC 53103	*Streptococcus mutans*	Biofilm formation	(141)
	GG		*gtfB*, *gtfC*, *gtfD* expression	(139)
L. acidophilus	-	*Klebsiella pneumoniae*	Replication inside biofilm	(136)
	DSM 20079	*Streptococcus mutans*	*gtfB* and *gtfC* expression	(142)
	-		*Gtf* and *LuxS* expression	(49)
A. kunkeei	-	*Pseudomonas aeruginosa*	Biofilm formation	(143)
L. plantarum	ATCC 10241	*Pseudomonas aeruginosa*	Biofilm formation and phagocytosis	(144)
	ATCC 14197	*Streptococcus mutans*	Biofilm formation	(138)
	299v DSM 9843		Biofilm formation	(141)
	-	*Streptococcus pyogenes*	Interleukin immune response	(145)
L. reuteri	DSM 20016	*Streptococcus mutans*	*gftB, gtfC* and *fft* expression	(47)
	ATCC 23272		Biofilm formation	(138)
	ATCC PTA 5289			(141)
	ATCC 55730			
	ATCC PTA-5289	*Streptococcus pyogenes*	Adhesion, hemolytic activity and *sag* expression	(140)
L. casei	4646	*Streptococcus mutans*	*luxS,* and *gftB, spaP, gbpB* expression	(146)
	ATCC 393		Biofilm formation	(138)

Supplementary table 9: Lactobacilli strains inhibiting HIV, *Neisseria gonorrhoeae, Candida albicans, Gardnerella veginalis, Trichomonas vaginalis, Prevotella bivia* and *Staphylococcus epidermidis* virulence factors

Lactobacilli	Strain	Pathogen	Effect	Reference
L. jensenii	-	HIV virus	Adhesion	(147)
		Neisseria gonorrhoeae	Adhesion	(148)
		Candida albicans	*ALS3, HWP1, ECE1* and *NRG1* expression	(149)
L. crispatus	CVT-05	UPEC	Adhesion	(150)
	-	*Gardnerella vaginalis*	*vly* and *sld* expression	(53)
	-	*Candida albicans*	*HWP1, ECE1*, *ALS3*, *BCR1*, *EFG1*, *TEC1* and *CPH1* expression	(151)
	ATCC 33820		Adhesion and interleukine immune response	(152)
	ATCC 33820		Adhesion and interleukine immune response	(153)
	B1-BC8		Adhesion	(154)
	-		*ALS3, HWP1, ECE1* and *NRG1* expression	(149)
L. acidophilus	T-13	*Staphylococcus epidermidis*	Adhesion	(155)
	ATCC 4356	*Candida albicans*	Adhesion and biofilm formation	(156)
	ATCC 4356		*ALS3, HWP1, ECE1* and *NRG1* expression	(149)
	ATCC 4356		Hyphal morphogenesis and biofilm	(157)
	T-13	UPEC	Adhesion	(155)
L. casei	ATCC 393	*Candida albicans*	Citotoxicity	(158)
	AMBR2		Hyphal morphogenesis	(56)
L. gasseri	ATCC 9857	*Trichomonas vaginalis*	Adhesion	(159)
	1	*Candida albicans*	Adhesion and biofilm formation	(156)
	-		Coaggregation	(160)
	-		*HWP1, ECE1*, *ALS3*, *BCR1*, *EFG1*, *TEC1* and *CPH1* expression	(151)
	BC9-BC14		Adhesion	(154)
	-		*ALS3, HWP1, ECE1* and *NRG1* expression	(161)
	KS120.1	*Prevotella bivia*	Adhesion	(162)
	ATCC 9857	UPEC		(150)
		Gardnerella vaginalis		
L. vaginalis	BC15-BC17	*Candida albicans*	Adhesion	(154)

L. fermentum	-	*Candida albicans*	Coaggregation	(160)
	-		*ALS3*, *HWP1*, *EFG1*, and *CPH1* expression	(163)
L. salivarius	ATCC 11741	*Candida albicans*	Citotoxicity	(158)
L. rhamnosus	ATCC 7469	*Candida albicans*	Hypae elongation	(158)
	CMP5351			
	GG ATCC 53103			
	GG ATCC 53103		Hyphal morphogenesis	(56)
	CMPG5351			
	CMPG5540			
	CMPG5357			
	CMPG10701			
	CMPG10706			
	GR-1 ATCC 5582			
	ATCC7469		Enzymatic activity and susceptibility to antifungals	(164)
	GG		Hyphal extention and adhesion	(165)
	GR-1		Interleukin immune response	(166)
	ATCC 9595		*BCR1, HWP1*, *ALS3* and *CPH1* expression	(167)
	-		*ALS3*, *HWP1*, *EFG1*, and *CPH1* expression	(163)
L. reuteri	ATCC 9595	*Candida albicans*	Adhesion and biofilm formation	(156)
	-		Coaggregation	(160)
	RC-14		Interleukin immune response	(166)
L. plantarum	-	*Candida albicans*	Coaggregation	(160)
	4B2	UPEC		(168)
		Streptococcus agalactiae		
		Gardnerella vaginalis		
L. paracasei	11	*Candida albicans*	Adhesion and biofilm formation	
	ATCC 11578		Citotoxicity	(158)
	ATCC 334		Hyphal morphogenesis	(56)
	-		*ALS3*, *HWP1*, *EFG1*, and *CPH1* expression	(163)
Unknown	-	*Candida albicans*	*HWp1*, *PLB2,* and *SAP1* expression	(169)
L. helveticus	KS300	*Garnerella vaginalis*	Adhesion	(170)
		UPEC	Adhesion, Invasion	
		Salmonella enterica serovar Typhimurium	Invasion	

Supplementary table 10: Lactobacilli strains inhibiting *Yersinia pseudotubercolosis, Yersinia enterocolitica, Serratia marcescens, Bacillus cereus, Enterococcus fecalis, Aggregatibacter actinomycetemcomitans,* and Rotavirus virulence factors

Lactobacilli	Strain	Pathogen	Effect on	Reference
L. acidophilus	LA 1	*Yersinia pseudotuberculosis*	Adhesion and invasion	(60)
	LB			(61)
	ATCC 4356	*Serratia marcescens*	Hemolytic activity and enzymatic expression	(171)
L. plantarum	C4	*Yersinia enterocolitica*	Immune system	(172)
	ATCC 8014	*Serratia marcescens*	Resistance to antibiotics and swarming mobility	(171)
	F14	*Bacillus cereus*	Hemolytic activity and enzymatic expression	(173)
	JX282192			
L. rhamnosus	-	*Enterococcus fecalis*	Immune system	(174)
L. reuteri	LMG P-27481	Rotavirus	Number of the copies	(116)
L. gasseri	OMZ525	*Aggregatibacter actinomycetemcomitans*	*LtxA* and *CdtB* expression	(57)
L. salivarius	OMZ520			

Chapter 2: Antibiotic resistance and virulence factors in lactobacilli: something to carefully consider

Andrea Colautti, Martina Arnoldi, Giuseppe Comi, Lucilla Iacumin*.

Department of Agricultural, Food, Environmental and Animal Science, University of Udine, via Sondrio 2/A, 33100 Udine, Italy

*Corresponding author:
Lucilla Iacumin,
Department of Food Science,
via Sondrio 2/A, 33100 Udine, Italy,
tel: +390432558126, Fax: +390432558130,
E-mail: lucilla.iacumin@uniud.it

Food Microbiology 103 (2022) 103934

Contents lists available at ScienceDirect

Food Microbiology

journal homepage: www.elsevier.com/locate/fm

Antibiotic resistance and virulence factors in lactobacilli: something to carefully consider

Andrea Colautti, Martina Arnoldi, Giuseppe Comi, Lucilla Iacumin *

Department of Agricultural, Food, Environmental and Animal Science, University of Udine, Via Sondrio 2/A, 33100, Udine, Italy

ARTICLE INFO

Keywords:
Lactobacilli
Antibiotic resistance
Virulence factor
Safety assessment
Resistance mechanisms
Genomics

ABSTRACT

Lactobacilli are a ubiquitous bacteria, that includes many species commonly found as part of the human microbiota, take part in the natural food fermentation processes, are used as probiotics, and in the food sector as starter cultures or bio-protectors. Their wide use is dictated by a long history of safe employ, which has allowed them to be classified as GRAS (General Recognized As Safe) microorganisms by the US Food and Drug Administration (FDA) and QPS (Qualified Presumption of Safety) by the European Food Safety Authority (EFSA, 2007; EFSA, 2021). Despite their classification as safe microorganisms, several studies show that some members of *Lactobacillus* genus can cause, especially in individuals with previous pathological conditions, problems such as bacteremia, endocarditis, and peritonitis. In other cases, the presence of virulence genes and antibiotic resistance, and its potential transfer to pathogenic microorganisms constitute a risk to be considered. Consequently, their safety status was sometimes questioned, and it is, therefore, essential to carry out appropriate assessments before their use for any purposes. The following review focuses on the state of the art of studies on genes that confer virulence factors, including antibiotic resistance, reported in the literature within the lactobacilli, defining their genetic basis and related functions.

1. Introduction

Antibiotic resistance (AR) is a natural bacterial mechanism. However, the inappropriate and generalized use of antibiotics has increased selective pressure resulting in the adaptation of bacteria to environmental changes and a related increase in resistance rates (Imperial and Ibana, 2016). Indeed, prolonged exposure to different concentrations of antibiotics can decrease the susceptibility of the bacterium, as demonstrated by Drago et al. (2011). Over the years, an increase in resistance rates has therefore been observed not only in pathogens but also in other microorganisms, including lactobacilli, indicating the bacteria previously belonging to the genus *Lactobacillus* given the recent reclassification into new 25 genera made by Zheng et al. (2020). This bacterial group is adapting to the environment by acquiring resistance genes from other resistant bacteria through a horizontal transfer mechanism (Imperial and Ibana, 2016; Lerner et al., 2019; Van Reenen and Dicks, 2011). This phenomenon is heightened by several factors, such as the increasingly selective stresses induced by clinical medication on the lactobacilli that colonize the human gastrointestinal tract (Ma et al., 2017) and from the wide use of antibiotics in the food chain (Willis, 2000), where often lactobacilli are intentionally added as starters. It is important to avoid that food becomes a promoter of new ARs or also a vector of them (Founou et al., 2016; Wang et al., 2006; McDermott et al., 2002; Van Reenen and Dicks, 2011). Therefore, in addition to limiting the use of antibiotics, and monitoring the presence of resistance factors in known pathogens, attention must be paid to the entire microbial population as recently pointed out also by EFSA (EFSA, 2007, EFSA, 2021), including lactobacilli,whose literature studies are lacking. Several authors showed that some lactobacilli can work as reserves of AR genes contributing to their potential transfer to pathogenic microorganisms, making antibiotics treatments ineffective (Egervärn et al., 2010; Gevers et al., 2003b; McDermott et al., 2002; Van Reenen and Dicks, 2011; Yang and Yu, 2019).

However, a specific AR profile in probiotics could be useful for concurrent use with antibiotics in the treatment of certain medical conditions (Anisimova and Yarullina, 2019; Hammad and Shimamoto, 2010). Therefore, it is necessary to highlight which are AR genes most commonly found in lactobacilli and the relative possibility of transfer to assess the safety of these bacteria. The crucial aspect of AR assessment is whether the latter is intrinsic or acquired. Intrinsic resistance is specific

* Corresponding author.
E-mail address: lucilla.iacumin@uniud.it (L. Iacumin).

https://doi.org/10.1016/j.fm.2021.103934
Received 25 February 2021; Received in revised form 25 October 2021; Accepted 4 November 2021
Available online 8 November 2021

2.1 Abstract

Lactobacilli are a ubiquitous bacteria, that includes many species commonly found as part of the human microbiota, take part in the natural food fermentation processes, are used as probiotics, and in the food sector as starter cultures or bio-protectors. Their wide use is dictated by a long history of safe employ, which has allowed them to be classified as GRAS (General Recognized As Safe) microorganisms by the US Food and Drug Administration (FDA) and QPS (Qualified Presumption of Safety) by the European Food Safety Authority (1, 2). Despite their classification as safe microorganisms, several studies show that some members of *Lactobacillus* genus can cause, especially in individuals with previous pathological conditions, problems such as bacteremia, endocarditis, and peritonitis. In other cases, the presence of virulence genes and antibiotic resistance, and its potential transfer to pathogenic microorganisms constitute a risk to be considered. Consequently, their safety status was sometimes questioned, and it is, therefore, essential to carry out appropriate assessments before their use for any purposes. The following review focuses on the state of the art of studies on genes that confer virulence factors, including antibiotic resistance, reported in the literature within the lactobacilli, defining their genetic basis and related functions.

Keywords

Lactobacilli, Antibiotic resistance, Virulence factor, Safety assessment, Resistance mechanisms, Genomics

2.2 Introduction

Antibiotic resistance (AR) is a natural bacterial mechanism. However, the inappropriate and generalized use of antibiotics has increased selective pressure resulting in the adaptation of bacteria to environmental changes and a related increase in resistance rates (Imperial and Ibana, 2016). Indeed, prolonged exposure to different concentrations of antibiotics can decrease the susceptibility of the bacterium, as demonstrated by Drago et al. (2011) (4). Over the years, an increase in resistance rates has therefore been observed not only in pathogens but also in other microorganisms, including lactobacilli, indicating the bacteria previously belonging to the genus *Lactobacillus* given the recent reclassification into new 25 genera made by Zheng et al. (2020) (5). This bacterial group is adapting to the environment by acquiring resistance genes from other resistant bacteria through a horizontal transfer mechanism (3,6,7). This phenomenon is heightened by several factors, such as the increasingly selective stresses induced by clinical medication on the lactobacilli that colonize the human gastrointestinal tract (8) and from the wide use of antibiotics in the food chain (9), where often lactobacilli are intentionally added as starters. It is important to avoid that food becomes a promoter of new ARs or also a vector of them (7, 10, 11). Therefore, in addition to limiting the use of antibiotics, and monitoring the presence of resistance factors in known pathogens, attention must be paid to the entire microbial population as recently pointed out also by EFSA (1,2), including lactobacilli, whose literature studies are lacking. Several authors showed that some lactobacilli can work as reserves of AR genes contributing to their potential transfer to pathogenic microorganisms, making antibiotics treatments ineffective (7, 12, 13, 14).

However, a specific AR profile in probiotics could be useful for concurrent use with antibiotics in the treatment of certain medical conditions (15, 16). Therefore, it is necessary to highlight which are AR genes most commonly found in lactobacilli and the relative possibility of transfer to assess the safety of these bacteria. The crucial aspect of AR assessment is whether the latter is intrinsic or acquired. Intrinsic resistance is specific for a bacterial species or genus. It has minimal possibility of horizontal transfer, while extrinsic resistance, which identifies a strain resistant to a specific antibiotic belonging to a typically sensitive taxonomic unit (17), has a high possibility of horizontal transfer. In the latter category, it is essential to distinguish resistance induced by chromosomal mutations from acquired genes, due to their higher transmission possibility for their possible collocation on mobile genetic elements, such as plasmids or transposons.

The horizontal transfer can occur through three mechanisms. By transformation, in which foreign genetic material is acquired from the extracellular environment (18); through the transduction mechanism, in which parts of bacterial DNA are included within a bacteriophage during replication, which subsequently infects another bacterial cell causing the transfer (7); or through the conjugation

process, in which the contact between cells induces the transfer of DNA (18). The latter mechanism, in which plasmids generally transport DNA, is the most commonly encountered in AR gene transfer (7), and is linked to the presence of mobile genetic elements such as plasmids, transposons, insertion sequences, bacteriophages (3,7,19).

AR genes can be successfully transferred using plasmids from lactobacilli to pathogenic or commensal bacterial strains and vice versa. In his work, Gevers et al. (2003) (20) highlighted the *in vitro* capacity of four strains of *Lactiplantibacillus plantarum*, two *Companilactobacillus alimentarius* and one *Latilactobacillus sakei* subsp. *sakei* to transfer by conjugation a tetracycline resistance gene to a strain of *Enterococcus faecalis* with a frequency ranging from 10 to 4 to 10–6 transconjugants per recipient. It has been found that transfer can also occur *in vivo* in an animal model (21, 22). Yang and Yu (2019) (14) demonstrated how tetracycline AR genes have been successfully transferred from *L. plantarum* and *Lactobacillus delbruekii* subsp. *bulgaricus*, isolated from yogurt to the pathogen *Listeria monocytogenes*. Successful plasmid acquisition and stability depend on various factors such as copy number, number of donors, specificity, and growth phase of the recipient organism (7, 23, 24). Transposons are DNA segments capable of interacting with other elements by recombination or transposition, repeatedly distributed on the chromosome in multiple copies, or associated with plasmids. The most common are the Tn917 and Tn916 families, generally associated with the transportation of the determinants for resistance to tetracycline and erythromycin (25). In the resistance evaluation, therefore, it is necessary to evaluate the presence of the relative genes and their chromosomal or plasmid positioning and to consider the possible presence of other mobile genetic elements. The analytical tests for AR include a phenotypic evaluation using various techniques such as E-test, disk diffusion test, and dilution, to determine the MIC value and to compare it with the appropriate species-specific threshold values (86). A strain is defined as resistant if the MIC value is greater than the threshold value, vice versa it is susceptible if the MIC value is lower (17). However, the phenotypic test alone is only useful as a preliminary test, as it does not differentiate between intrinsic and extrinsic resistance. Furthermore, the MIC cut-off values, especially in lactobacilli, are not standardized, and this can lead to ambiguous results in considering a microorganism resistant or not. Also, a negative phenotypic result does not guarantee the absence of transferable resistance genes (15, 26, 27). It is, therefore, useful to combine molecular analysis and genetic sequencing to identify the possible presence of genes related to phenotypic resistance and avoid false assumptions (28). However, the presence of a gene in the genome does not always generate resistance: the possible presence of stop codons, insertions, or deletions could make the gene non-functional. In any case, even if the considered microorganisms are not able to express their relative resistance, the transfer to commensal and pathogenic bacteria cannot be excluded (29).

In addition to the antibiotic resistance factors, there are other important genes to consider. In some cases, these organisms have been associated with diseases such as peritonitis, infectious endocarditis, bacteraemia, and urinary tract infections (6, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46) mainly in immunocompromised patients. The most commonly involved and documented species consist of *Lacticaseibacillus rhamnosus*, *Lacticaseibacillus paracasei,* and *Lacticaseibacillus casei* (30, 37, 41, 47, 48), corresponding to the most frequently used species as probiotic cultures.
Generally, lactobacilli isolated from infected clinical samples coincide with lactobacilli of the intestinal microbiota, but cases have been reported in which clinical isolates phylogenetically correspond to lactobacilli administered in the form of probiotics (30, 37). Yelin et al. (2019) (49) highlighted a greater risk of contracting *Lactobacillus* bacteraemia in intensive care unit (ICU) patients treated with probiotics than in untreated patients, noting a phylogenetic similarity between lactobacilli isolated from infected blood and the administered probiotic *L. rhamnosus GG*, underlining its ability to reach the bloodstream. Therefore, the identification of potential virulence traits among lactobacilli is useful to evaluate the safety of these bacteria before their usage in the food industry and as human probiotics. The virulence factors mainly associated with lactobacilli consist in the ability of some strains to produce specific enzymes as glycosidases and arylamidase proteases (50), and proteins capable of binding fibrinogen, collagen, and fibronectin, inducing bacterial migration determining the subsequent evasion from host defence mechanisms and the possibility of platelet aggregation (33, 48, 51, 52). Some lactobacilli have also shown the presence of virulence genes generally associated with pathogenic microorganisms, such as *Enterococcus* spp. and *Staphylococcus* spp. due to the acquisition by horizontal transfer.
The following review deal with genes that confer virulence factors, including antibiotic resistance, found so far in the literature within lactobacilli, defining their genetic basis and related functions.

2.3 Antibiotic resistance genes found in lactobacilli

Numerous studies highlighted how lactobacilli can be considered reserves of AR genes. The resistance genes to various antibiotics reported in literature for these bacteria were presented below, defining the resistance mechanism and the possible transfer to pathogens.

2.3.1 Tetracycline

One of the most evaluated resistances was that referred to tetracycline. Tetracycline is a broad-spectrum antibiotic that exhibits activity against a wide range of Gram-positive and Gram-negative bacteria. Its role is the inhibition of protein synthesis by binding the conserved 16 S rRNA sequences of the 30 S subunit of the ribosome to prevent attachment to A-site by t-RNA (53).

Resistance is related to the presence of two gene groups. The first group is responsible for the production of membrane-associated proteins capable of mediating the antibiotic efflux outside the cell, reducing its intracellular concentration. Genes *tet*(Z), *tet*(K), *tet*(L), and *tcr*3 belong to this category. A second group is composed of genes that encode cytoplasmic proteins able to protect ribosomes from antibiotic attack: *tet*(M), *tet*(O), *tet*(S), *tet*(W), *tet* (Q), *tet*(T), and *otr*(A) (54). A large number of lactobacilli carrying one or more genes related to tetracycline resistance have been reported (Table 1). *tet*(M) gene represented the most widespread determinant. It was found in numerous strains of *L. plantarum* isolated from animals and products of animal origin (13, 21, 24, 26, 55), silage (12), and green tea (28). Different strains of *Limosilactobacillus reuteri*, *Ligilactobacillus salivarius,* and *L. sakei* showed the same result (Table 1). Todorov et al. (2019) (56) showed the ribosome protection gene presence in three tested strains of *Latilactobacillus curvatus* isolated from smoked salmon, as well as Yang and Yu (2019) (14) in three strains of *L. delbrueckii* subsp. *bulgaricus* from yogurt samples. These data underlined the possible spread of this determinant in numerous food products, creating concern about the possible transfer within the gastrointestinal tract. *tet*(M) is generally associated with transposons of the Tn916 - Tn1545/Tn917 family identified by the presence of the *int* integrase and *xis* excision genes that mediate its transfer (21, 22, 25, 29, 57, 58). Furthermore, several studies highlight the presence of this gene on plasmids, contributing to the possibility of its spread also to pathogenic microorganisms (12, 13, 26), such as *Listeria monocytogenes* (14), while others noted a chromosomal positioning (59).

Table 1 - Tetracycline resistance genes found in *Lactobacillus* spp.

Ribosomal protection genes

Gene	Species	Reference
tet(W)	*L. amylovorus*	(Chang et al., 2011)
	L. kefiri	
	L. parabuchneri	
	L. ruminis	
	L. salivarius	(Chang et al., 2011; Thumu and Halami, 2012)
	L. helveticus	(Guo et al., 2017)
	L. paracasei	(Huys et al., 2008)
	L. reuteri	(Chang et al., 2011; Egervarn et al., 2010; Egervarn et al., 2009; Kastner et al., 2006; Thumu and Halami, 2012)
	L. sakei	(Zonenschain et al., 2009)
	L. curvatus	
	L. plantarum	(Chang et al., 2011; Thumu and Halami, 2012; Zonenschain et al., 2009)
	L. rhamnosus	(Chang et al., 2011; Thumu and Halami, 2012; Zhang e Zhang, 2019; Zonenschain et al., 2009)
	L. fermentum	(Chang et al., 2011; Zhang and Zhang, 2019)
	L. delbrueckii	(Campedelli et al., 2019)

	L. reuteri - vaccinostercus	
tet(O)	*L. curvatus*	(Todorov et al., 2019)
	L. delbrueckii	
	L. plantarum	(Arellano et al., 2019; Zhang and Zhang, 2019)
	L. rhamnosus	(Zhang and Zhang, 2019)
	L. salivarius	(Aquilanti et al., 2007; Thumu and Halami, 2012)
tet(Q)	*L. brevis*	(Campedelli et al., 2019)
	L. salivarius	(Chang et al., 2011)
tet(S)	*L. plantarum*	(Arellano et al., 2019; Yang e Yu, 2019; Zonenschain et al., 2009)
	L. paraplantarum	(Ouoba et al., 2008)
	C. alimentarius	(Campedelli et al., 2019)
tet(W/O)	*L. johnsonii*	(Ammor et al., 2008a)
tet(M)	*L. amylovorus*	(Chang et al., 2011)
	L. kefiri	
	L. parabuchneri	
	L. pentosus	(Preethi et al., 2017)
	L. paracasei	(Devirgiliis et al., 2009; Huys et al., 2008; Zonenschain et al., 2009)
	L. delbrueckii subsp. *bulgaricus*	(Yang and Yu, 2019)
	L. salivarius	(Aquilanti et al., 2007; Cauwerts et al., 2006a; Chang et al., 2011; Preethi et al., 2017; Thumu e Halami, 2012)
	L. brevis	(Zonenschain et al., 2009)
	L. rhamnosus	
	L. sakei	(Chang et al., 2011; Gevers et al., 2003a; Zonenschain et al., 2009)
	L. curvatus	(Todorov et al., 2019; Zonenschain et al., 2009)
	L. reuteri	(Aquilanti et al., 2007; Chang et al., 2011; Zonenschain et al., 2009)
	L. plantarum	(Aquilanti et al., 2007; Arellano et al., 2019; Chang et al., 2011; Egervarn et al., 2009; Gevers et al., 2003a; Preethi et al., 2017; Zonenschain et al., 2009)
	L. casei - maniotivorans	(Campedelli et al., 2019)
	L. delbrueckii	
	L. reuteri - vaccinostercus	
*otr*A	*L. fermentum*	(Zhang and Zhang, 2019)
tet(T)	*L. fermentum*	(Zhang and Zhang, 2019)
	L. rhamnosus	
	L. plantarum	
tet(O/W/32/O/W/O)	*L. johnsonii*	(van Hoek et al., 2008)

Efflux proteins genes		
Gene	**Species**	**Reference**
tet(K)	*L. amylovorus*	(Chang et al., 2011)
	L. buchneri	(Anisimova e Yarullina, 2019)
	L. curvatus	(de Castilho et al., 2019; Todorov et al., 2019)
	L. delbrueckii	(Todorov et al., 2019)

	L. fermentum	(Chang et al., 2011; Thumu and Halami, 2012)
	L. plantarum	(Aquilanti et al., 2007; Arellano et al., 2019; Chang et al., 2011; Todorov et al., 2017)
	L. reuteri	(Aquilanti et al., 2007)
	L. ruminis	(Chang et al., 2011)
	L. salivarius	(Aquilanti et al., 2007; Chang et al., 2011)
tet(PB)	*L. fermentum*	(Zhang and Zhang, 2019)
	L. plantarum	
tet(Z)	*L. reuteri*	(Cauwerts et al., 2006b)
tet(L)	*L. curvatus*	(de Castilho et al., 2019)
	L. plantarum	(Anisimova and Yarullina, 2019; Arellano et al., 2019; Chang et al., 2011; Preethi et al., 2017; Thumu and Halami, 2012)
	L. amylovorus	(Chang et al., 2011)
	L. kefiri	
	L. parabuchneri	
	L. reuteri	
	L. ruminis	
	L. sakei	
	L. fermentum	(Thumu and Halami, 2012)
	L. salivarius	(Chang et al., 2011; Preethi et al., 2017; Thumu and Halami, 2012)
tcr3	*L. fermentum*	(Zhang and Zhang, 2019)
	L. rhamnosus	
	L. plantarum	

Other widespread genes within lactobacilli resulted in the *tet*(W) gene, which encodes a ribosome protection protein, and *tet*(K) and *tet*(L) genes, responsible for the antibiotic efflux (Table 1). Chang et al. (2011) (55) showed that on 146 strains of 11 species of lactobacilli isolated from pig intestine, 82.0% of tetracycline-resistant (TETR) strains had *tet*(W) gene, 22.5% *tet*(M) gene, 14.4% *tet*(L) gene and 8.1% *tet*(K) gene. The same authors were also the first to find the presence of the *tet*(Q) gene in a strain of *L. salivarius*. Although in some cases *tet*(W) gene has been found on non-conjugative plasmids and the possibility of transfer was not prooved (12, 27, 60), there are pieces of evidence of its presence on plasmids, and therefore the possibility of transfer cannot be excluded. Thumu and Halami (2019) (22) highlighted the presence of a plasmid containing *tet*(W), *tet*(M), and *tet*(L) genes that were transferred from a strain of *L. salivarius* to the recipient *Enterococcus faecalis* JH2-2, both *in vitro* and *in vivo*. It is therefore possible to find multiple *tet* genes in the same microorganism, with the same (ribosomal protection or efflux) or different action (efflux and ribosomal protection) (22, 24, 61). Other genes that confer resistance to tetracycline were frequently observed in strains belonging to *L. plantarum* species such as *tet*(S), *tet*(T), and *tet*(O) (14, 23, 24, 28, 62). *tcr*3 genes found mainly in *Streptomyces* spp. (54) and *otr*A, for resistance to oxytetracycline, were recently found in strains belonging to the species *Limosilactobacillus fermentum*, *L. rhamnosus,* and *L. plantarum* (62). Two tetracycline resistance mosaic genes were also observed, resulting from

interclass recombination within the coding regions of the *tet*W and *tet*O genes, such as *tet* (W/O) (63) and *tet* (O/W/32/O/W/O) in a strain of *Lactobacillus johnsonii* isolated from human feces (64) (Table 1).

There are numerous data in the literature regarding tetracycline resistance for L. *plantarum* (Table 1). Most of the studies found a *tet* + profile in phenotypically resistant strains (Table 2). However, in a study conducted by Anisimova and Yarullina (2019) (15), on 12 strains of *L. plantarum* tested, all showed susceptibility to the antibiotic, 25% (3/12) of which, however, had a positive response for the gene *tet*(L). Arellano et al. (2020) (28) presented a similar result, in which 11 out of 18 tested strains were susceptible to tetracycline, despite having different resistance genes. These results highlighted how a negative phenotypic result cannot exclude the presence of potentially transmissible genes. Comparing, in Table 2, the number of strains of *L. plantarum* resistant to tetracycline to the corresponding number of resistance genes found, it was noted that the percentage of resistance genes is relatively low (21, 24, 26, 65). This finding should not be confused with the possibility of intrinsic resistance, as not all possible genes responsible for resistance were always evaluated. For example, Campedelli et al. (2019) (65) considered only *tet*(W), *tet*(M), *tet*(L), *tet*(P), *tet*(S) e *tet*(Q) determinants.

The ranges of MIC values (Table 2) for *L. plantarum* were found to be relatively broad in all the considered studies, with values higher than 512 μg/mL (12, 21, 24, 55). Other well-documented tetracycline-resistant lactobacilli species were *L. reuteri*, *L. sakei*, and *L. salivarius*. As reported in Table 3, different studies highlighted MIC values even above 32 μg/mL. Chang et al. (2011) (55) observed resistance with MIC values ranging from 16 to 1024 μg/mL for all 60 strains of *L. reuteri* tested, in which the predominantly detected gene was found to be *tet*(W) (43/60 strains tested), and for 100% of the *L. sakei* (6/6) and *L. salivarius* (17/17) strains examined. Concerning these species, phenotypic resistance is not always correlated to the presence of *tet* genes, probably because all the possible determinants able to induce resistance to tetracycline are not always considered 24, 55, 65, 66).

Table 2 – Resistance to tetracyclin in *L. plantarum*

Species	Origin	Tested Strains	Phenotype (n° strains /n°tested strains)	MIC range (µg/mL)	Cut-off (µg/mL)	Genotype	n° strains with gene /n°tested strains	Reference
L. plantarum	salami	8	8/8 R	>256 [d]	ND	*tet*(M)	8/8	(Gevers et al., 2003a)
	meat products	11	3/3 R**	16 -> 64[c]	16[e]	*tet*(M)	2/11	(Aquilanti et al., 2007)
						tet(K)	2/11	
	salami	12	11/12 R	16 - 512[c]	32[f]	*tet*(M)	5/11	(Zonenschain et al., 2009)
						tet(W)	4/11	
						tet(S)	1/11	
	silage	2	2/2 R	>256[d]	ND	tet(M)	2/2	(Egervarn et al., 2009)
	swine colon	11	11/11 R	32 - 512[c]	32[g]	*tet*(M)	4/11	(Chang et al., 2011)
						tet(K)	1/11	
						tet(L)	1/11	
	ice cream	1	1/1 R	128[c]	32[g]	*tet*(W), *tet*(L)	1/1	(Thumu and Halami, 2012)
	poultry feces and intestines	10	10/10 R	4 - 1024[c]	ND	*tet*(L)	4/10	(Preethi et al., 2017)
						tet(M)	5/10	
	salami	1	ND	*ND*	ND	*tet*(K), *tet*(O)	1/1	(Todorov et al., 2017)
	silage	12	12/12 S	≥19mm[b]	≤14 mm	*tet*(L)	3/12	(Anisimova and Yarullina, 2019)
	green tea, fermented products, insects	18	7/18 R	≤8 ->64[a]	32[h]	*tet*(M)	2/2 *	(Arellano et al., 2019)
						tet(K)	2/2 *	
						tet(L)	2/2 *	
						tet(O)	2/2 *	
						tet(S)	2/2 *	
	fermented dairy products	6	1/6 R	1 - 32[c]	32[g]	*tet*(S)	1/6	(Yang and Yu, 2019)
	caries	7	1/7 R	16 - 64[a]	32[g]	*tet*(T).t*et*(L), *tet*(O), *tcr*3, *tet*PB	1/7	(Zhang and Zhang, 2019)
	Pickles, beer contaminant, fermented vegetables	10	7/10 R	4 – 64[c]	ND	-	0/10	(Campedelli et al., 2019)

R: resistant, S: susceptible, ND: not defined

a: MIC evaluated with agar dilution method (µg/mL)

b: MIC evaluated with disk diffusion method (mm)

c: MIC evaluated with broth microdilution method (µg/mL)

d: MIC evaluated with E-Test (µg/mL)

e: Cut-off MIC value for tetracycline defined by CLSI (2008)

f: Cut-off MIC value for tetracycline defined by EFSA (2005)

g: Cut-off MIC value for tetracycline defined by EFSA (2008)

h: Cut-off MIC value for tetracycline defined by EFSA (2012)

*: genotypic analysis performed only on 2 phenotypically susceptible strains

**: phenotypic analysis carried out on the 3 strains with the related gene

Table 3 - Distribution of tetracycline resistance in *L. reuteri, L. sakei* and *L. salivarius*

Species	Origin	Tested Strains	Phenotype (n° strains /n°tested strains)	MIC *range* (µg/mL)	*Cut-off* (µg/mL)	Genotype	n° strains with gene /n°tested strains	Reference
L. reuteri	different origins	32	28/32 R	*4 - >256*[d]	ND	*tet*(W)	24/28	(Egervarn et al., 2009)
	swine colon	60	60/60 R	*16 - 1024*[c]	16[g]	*tet*(M) *tet*(W) *tet*(L)	2/60 43/60 4/60	(Chang et al., 2011)
	salami	2	2/2 R	*256- >512*[c]	16[g]	*tet*(W)	1/2	(Thumu e Halami, 2012)
	meat products	3	2/2 R**	*32 - 64*[c]	16[e]	*tet*(M) *tet*(K)	1/3 1/3	(Aquilanti et al., 2007)
	salami	1	1/1 R	*512*[c]	8[f]	tet(M)	1/1	(Zonenschain et al., 2009)
	cloaca of broiler poultry	8	8/8 R	≥64[a]	16	*tet*(L) *tet*(W) tet(L)+ *tet*(K) tet(Z)	1/8 3/8 1/8 1/8	(Cauwerts et al., 2006a)
	animal feces, vagina, cheese, sour dough, human and pig intestines, human saliva	18	12/18 R	1 – 128[c]	ND	*tet*(L) *tet*(M) *tet*(W)	2/12 2/12 1/12	(Campedelli et el., 2019)
L. sakei	salami	24	17/24 R	*2 - 512*[c]	8[f]	*tet*(M) *tet*(W)	11/17 1/17	(Zonenschain et al., 2009)
	salami	10	10/10 R	*32 - >256*[d]	ND	*tet*(M)	10/10	(Gevers et al., 2003a)
	swine colon	6	6/6 R	*32 - 256*[c]	8[g]	*tet*(M) *tet*(L)	4/6 2/6	(Chang et al., 2011)
	fermented meat products, sake starters, rice noodles, silage, milk	5	1/5	*0,5 – 16*[c]	ND	-	0/1	(Campedelli et al., 2019)
L. salivarius	cloaca of broiler poultry	31	24/31 R	*2 -> 64*[a]	16	*tet*(M) *tet*(L)+ *tet*(M)	9/24 14/24	(Cauwerts et al., 2006a)
	swine colon	17	17/17 R	*16 - 512*[c]	8[g]	*tet*(M) *tet*(W) *tet*(K) *tet*(L) *tet*(Q)	5/17 8/17 1/17 2/17 1/17	(Chang et al., 2011)
	salami	3	3/3 R	*256*[c]	16[e]	*tet*(M) *tet*(W) *tet*(O) *tet*(L)	3/3 3/3 1/3 1/3	(Thumu e Halami, 2012)
	meat products	6	6/6 R	*>64*[c]	16[e]	*tet*(M) *tet*(O) *tet*(K)	4/6 1/6 2/6	(Aquilanti et al., 2007)
	intestines and poultry meat, slaughter water	3	3/3 R	*4 - 128*[c]	ND	*tet*(M)	2/3	(Preethi et al., 2017)
	cider, must, animal faeces, rat and chicken intestines, saliva, slurry, vacuum-packed meat	27	5/27	*0,5 – 128*[c]	ND	*tet*(P)	1/5	(Campedelli et al., 2019)

R: resistant; S: susceptible; ND: not defined; **: phenotypic analysis carried out on the 2 strains with the related gene

a: MIC evaluated with agar dilution method (µg/mL)

b: MIC evaluated with disk diffusion method (mm)

c: MIC evaluated with broth microdilution method (µg/mL)

d: MIC evaluated with E-Test (µg/mL)

e: Cut-off MIC value for tetracycline defined by CLSI (2008)

f: Cut-off MIC value for tetracycline defined by EFSA (2005)

g: Cut-off MIC value for tetracycline defined by EFSA (2008)

2.3.2 MLS - macrolides, lincosamides, streptogramins

Another widespread gene in lactobacilli reported in the literature was erm(B) (Table 4), which is linked to resistance to erythromycin, a macrolide belonging to the MLS group of antibiotics. The antibiotics present within the group (erythromycin, clindamycin, lincomycin, and streptogramin A) have a different chemical structure but share the same protein synthesis inhibition action. Resistance can occur through three mechanisms: methylation of the target site of the antibiotic, efflux, and inactivation. The methylation mechanism confers a broad spectrum of resistance to macrolides and lincosamides, while antibiotic efflux and inactivation activities are targeted only to certain antibiotics or classes of antibiotics (25, 67).

Table 4 - Antibiotic resistance genes of the MLS group found in Lactobacillus spp. with methylation action of the target site of the antibiotic

Gene	Species	Reference
erm(B)	*L. helveticus*	(Guo et al., 2017)
	L. pentosus	(Preethi et al., 2017)
	L. casei	(de Souza et al., 2019; Guo et al., 2017)
	L. crispatus	(Aquilanti et al., 2007)
	L. johnsonii	
	L. sakei	(Zonenschain et al., 2009)
	L. curvatus	
	L. paracasei	(Huys et al., 2008; Zonenschain et al., 2009)
	L. brevis	(Zonenschain et al., 2009)
	L. rhamnosus	
	L. fermentum	(de Souza et al., 2019; Thumu and Halami, 2012; Todorov et al., 2019)
	L. salivarius	(Aquilanti et al., 2007; Hummel et al., 2007; Preethi et al., 2017; Thumu and Halami, 2012)
	L. plantarum	(Anisimova and Yarullina, 2019; Aquilanti et al., 2007; Arellano et al., 2019; Feld et al., 2009; Guo et al., 2017; Preethi et al., 2017; Thumu and Halami, 2012; Todorov et al., 2017; Zonenschain et al., 2009)
	L. reuteri	(Aquilanti et al., 2007; Egervarn et al., 2009; Ouoba et al., 2008; Thumu and Halami, 2012; Zonenschain et al., 2009)
	L. delbrueckii	(Campedelli et al., 2019)
	L. reuteri - vaccinostercus	
erm(A)	*L. curvatus*	(Todorov et al., 2019)
	L. delbrueckii	
	L. fermentum	
	L. plantarum	(Arellano et al., 2019)
erm(C)	*L. plantarum*	(Arellano et al., 2019; Todorov et al., 2017; Zonenschain et al., 2009)
	L. brevis	(Aquilanti et al., 2007)
	L. johnsonii	
	L. reuteri	(Egervarn et al., 2009)
	L. casei	(de Souza et al., 2019)
	L. fermentum	(Todorov et al., 2019)
	L. delbrueckii	
	L. curvatus	(de Castilho et al., 2019; Todorov et al., 2019)
erm(LF)	*L. fermentum*	(Gfeller et al., 2003)
erm(T)	*L. reuteri*	(Egervarn et al., 2009)

Erythromycin resistance is mediated by genes of the *erm* (*Erythromycin Ribosome Metylase*) class encoding a ribosomal adenine-N6-methyltransferase. These genes can hinder the action of the antibiotic by methylating the 23 S rRNA peptidyltransferase center, thus preventing the antibiotic from attacking the ribosome 50 S subunit (53). As can be seen in Table 4, which represents the diffusion of erythromycin resistance genes in *Lactobacillus* spp, the most commonly found classes in this group are *erm*(A), *erm*(B), and *erm*(C), detected in several strains of *L. plantarum*, *L. casei*, *L. reuteri*, *L. rhamnosus* and *L. fermentum* (12,, 15, 21, 26, 28, 29, 56, 58, 60, 68, 69). It can be noted that, even in the case of erythromycin, several studies have observed numerous strains of *L. plantarum* endowed with at least one of the genes listed above (Table 4). In this case, however, unlike the results obtained for resistance to tetracycline, most of the strains tested in the various studies were found to be susceptible to the antibiotic (Table 5). Generally, lactobacilli are susceptible to antibiotics that inhibit protein synthesis such as erythromycin (14, 58), but the increasing selective pressure has made it possible to adapt and acquire resistance-related genes. Guo et al. (2017) (60)found that a total of 33 lactobacilli tested (11 *Lactobacillus helveticus*, 11 *L. casei*, 11 *L. plantarum*) were sensitive to erythromycin (MIC range 0.016–1 μg/mL), but only 6 strains of *L. helveticus*, 1 strain of *L. casei* and 1 strain of *L. plantarum* reported the presence of the *erm*(B) gene. A comparable result was obtained by Anisimova and Yarullina (2019) (15), who showed that on 20 lactobacilli strains tested with negative phenotype for resistance, 4 strains of *L. plantarum* carried the corresponding resistance genes, highlighting the possible presence of acquired silent genes. In contrast, Thumu and Halami (2012) (22), considering the used cut-off of 1 μg/mL, observed high resistance for all 10 lactobacilli strains tested (MIC range 8–512 μg/mL), each of which carried at least one resistance gene. As reported in Table 5, MIC values for erythromycin in the most studied species (*L. salivarius, L. reuteri*, and *L. plantarum*) were quite different, at both inter-species and intra-species levels. The presence of silent genes with the relative negative phenotypic outcome should cause concern for possible transfer. Although in some cases *erm*(B) gene was found to be positioned at the chromosomal level, reducing the possibility of transfer (69), in others a plasmid positioning was observed, which increase the transfer probability (12, 13). Indeed, the transmission of erythromycin resistance from strains of *L. reuteri*, *L. plantarum,* and *L. salivarius* to *Enterococcus faecalis JH2-2* has been demonstrated in *vitro* conjugation experiments (22, 23). Feld et al. (2009) (70) showed the ability of a strain of *L. plantarum* to transfer, through the plasmid pLFE1, the *erm*(B) gene to *L. rhamnosus*, *Lactococcus lactis*, *Listeria innocua*, to the opportunistic pathogen *Enterococcus faecalis* and the pathogen *Listeria monocytogenes*.

Table 5 - Distribution of resistance to erythromycin in *L. reuteri, L. salivarius* and *L. plantarum*

Species	Origin	Tested Strains	Phenotype (n° strains /n°tested strains)	MIC range (µg/mL)	Cut-off (µg/mL)	Genotype	n° strains with gene /n°tested strains	Reference
L. reuteri	meat products	3	1/1 S**	*0,25*[c]	1[e]	*erm*(B)	1/3	(Aquilanti et al., 2007)
	swine feces	1	1/1 R	*>32*[c]	8[f]	*erm*(B)	1/1	(Ouoba et al., 2008)
	different origins	32	6/32 R	*0,25 - >256*[d]	ND	*erm*(B)	4/6	(Egervarn et al., 2009)
						erm(C)	1/6	
						erm(T)	1/6	
	salami	1	1/1 R	*512*[c]	4[g]	*erm*(B)	1/1	(Zonenschain et al., 2009)
	salami	2	2/2 R	*256 - 512*[c]	1[h]	*erm(*B)	2/2	(Thumu and Halami, 2012)
	animal feces, vagina, cheese, sourdough, swine and human intestine, human saliva	18	3/18	0,016 – 16[c]	ND	*erm*(B)	1/3	(Campedelli et al., 2019)
L. salivarius	meat products	6	2/3 R**	*0,25 - >32*[c]	1[e]	*erm*(B)	3/6	(Aquilanti et al., 2007)
	salami	3	3/3 R	*64 - 256*[c]	≥1[l]	*erm(*B)	3/3	(Thumu and Halami, 2012)
	poultry meat and intestine, slaughter water	3	3/3 R	*4 - 128*[c]	ND	*erm(*B)	2/3	(Preethi et al., 2017)
						msrA/B	1/3	
	Cider, must, animal faeces, rat and chicken intestines, saliva, slurry, vacuum-packed meat	27	5/27	*0,06 – 16*[c]	ND	-	0/5	(Campedelli et al., 2019)
L. plantarum	green tea, fermented foods, insects	18	1/18 R	0,5 - 2[a]	1[g]	*erm*(A)	2/2 *	(Arellano et al., 2019)
						erm(B)	2/2 *	
						erm(C)	2/2 *	
	silage	12	11/12 S 1/12 MS	*>14 mm*[b]	≤13 mm	*erm*(B)	3/12	(Anisimova e Yarullina, 2019)
						mef(A)	1/12	
	meat products	11	6/6 S **	*0,125 - 0,25*[c]	1[e]	*erm*(B)	6/11	(Aquilanti et al., 2007)
	salami	12	6/12 R	*0,25 - 512*[c]	4[g]	*erm*(B)	3/6	(Zonenschain et al., 2009)
						erm(C)	2/6	
	ice cream	1	1/1 R	*32*[c]	1[h]	*erm*(B)	1/1	(Thumu and Halami, 2012)
	dairy products	11	11/11 S	*0,0625 - 0,25*[c]	1[i]	*erm*(B)	1/11	(Guo et al., 2017)
	salami	1	ND	ND	ND	*erm*(B), *erm*(C)	1/1	(Todorov et al., 2017)
	poultry meat and intestine	10	10/10 R	*4 - 1024*[c]	ND	*erm*(B)	8/10	(Preethi et al., 2017)
						*msr*A/B	3/10	
						*msr*C	2/10	
	pickles, beer contaminant, fermented vegetables	10	2/10	*0,25 – 16*[c]	ND	-	0/2	(Campedelli et al., 2019)

R: resistant, S: susceptible, MS: moderatlely susceptible ND: not defined

a: MIC evaluated with agar dilution method (µg/mL)

b: MIC evaluated with disk diffusion method (mm)

c: MIC evaluated with broth microdilution method (µg/mL)

d: MIC evaluated with E-Test (µg/mL)

e: Cut-off MIC value for tetracycline defined by Danielsen and Wind (2003)

f: Cut-off MIC value for tetracycline defined by Ouoba et al., 2008

g: Cut-off MIC value for tetracycline defined by EFSA (2005)

h: Cut-off MIC value for tetracycline defined by EFSA (2008)

i: Cut-off MIC value for tetracycline defined by EFSA (2012)

l: Cut-off MIC value for tetracycline defined by CLSI (2008)

*: genotypic analysis performed only on 2 phenotypically susceptible strains

**: phenotypic analysis carried out on the strains showing the related gene

Thumu and Halami (2019) (22) demonstrated how the transfer of a plasmid containing the *erm*(B) gene in association with the *tet*(M), *tet*(W), and *tet*(L) genes could also occur *in vivo*. There is also the possibility of finding genes for resistance to erythromycin associated with mobile elements, such as the transposon Tn917, capable of carrying the *erm*(B) gene, or transposons of the Tn916 family able to carry the determinants for resistance to erythromycin associated to the ones to tetracycline (25).

Although the *erm*(A), *erm*(B), and *erm*(C) genes were the most found and documented determinants in *Lactobacillus* spp, the presence of the *erm*(T) gene was detected in a strain of *L. reuteri* isolated from poultry intestine (71) and *erm* (LF) gene in a transferable plasmid of a *L. fermentum* strain (59). Another mechanism through which resistance to erythromycin can occur is by reducing the intracellular concentration of the antibiotic, thanks to the presence of efflux pumps encoded by the *mef* genes, for example, *mef*(A), observed in *L. plantarum* and *L. salivarius* subsp. *salivarius* (15, 72), and *mef*(B) and *mef*(E) found in *L. casei and L. delbrueckii* (65). Other antibiotic efflux-related genes are *msr* (A/B) found in *L. plantraum* and *L. salivarius,* and the enterococcal gene *msr*(C) discovered in *L. fermentum* and *L. plantarum* (21, 125) (Table 6). This last gene confers resistance to both macrolides and group B streptogramins (25). Campedelli et al. (2019) (65) mentioned also the *lsa* gene, related to the efflux of the lincosamide clinamycin, which was found in 60 lactobacilli strains, 13 of which resistant to the corresponding antibiotic. The third mechanism of resistance consists in the inactivation of the antibiotic, mediated by the *lnu*(A) gene encoding a transferase capable of inhibiting the lincosamides action. As showed in Table 7, this gene was observed in strains of *L. reuteri* found to be resistant to clindamycin and lincomycin (27, 72). Kastner et al. (2006) (27) noted a 96% similarity between *L. reuteri* SD 2112 *lnu*(A) and *Staphylococcus haemolyticus lin*(A) gene sequence. An acetyltransferase encoded by the *vat*(E) gene, able to inactivate group A streptogramins (e.g. dalfopristin), was observed in strains of *L. curvatus* and *L. fermentum* (56) and in ROT1 isolated from cheese associated with *erm* (LF) gene, placed on the pLME300 plasmid, able to confer high resistance to dalfopristin and erythromycin (55) (Table 7).

Table 6- Antibiotic resistance genes of the MLS group found in *Lactobacillus* spp. with efflux action

Gene	Antibiotic	Species	Reference
mef(A)	Erythromycin	*L. salivarius* subsp. *salivarius*	(Cauwerts et al., 2006b)
		L. plantarum	(Anisimova and Yarullina, 2019)
mef(B)	Erythromycin	*L. casei - maniotivorans*	Campedelli et al., (2019)
mef(E)	Erythromycin	*L. delbrueckii*	
msr(A/B)	Erythromycin	*L. plantarum*	(Preethi et al., 2017)
		L. salivarius	
msr(C)	Erythromycin + Streptogramin B	*L. fermentum*	(Thumu e Halami, 2012)
		L. plantarum	(Preethi et al., 2017)
lsa	Clindamycin	*L. delbrueckii*	(Campedelli et al., 2019)
		L. reuteri	
		L. sakei	
		L. brevis	
		L. buchneri	
		L. casei	
		L. collinoides	
		L. coryniformis	
		C. alimentarius	

Table 7 - Antibiotic resistance genes of the MLS group found in *Lactobacillus* spp. with an antibiotic inactivation action

Gene	Antibiotic	Species	Reference
lnu(A)	clindamycin, lincomycin	*L. reuteri*	(Cauwerts et al., 2006b; Kastner et al., 2006)
*vat(*E)	streptogramin A, dalfopristin	*L. curvatus*	(Todorov et al., 2019)
		L. fermentum	(Gfeller et al., 2003; Todorov et al., 2019)

2.3.3 Aminoglycosides

Aminoglycosides are antibiotics responsible for interfering with protein synthesis by binding to the acceptor (A) site placed on the 16 S rRNA of the 30 S ribosomal subunit, resulting in translation block (53, 73). One resistance mechanism is the inactivation of the antibiotic mediated by intracellular enzymes able to modify its structure, inducing the reduction of affinity for A site, and consequently preventing the binding to the 30 S ribosomal subunit. Within this group, there are 4 classes of enzymes: aminoglycosides phosphotransferases (APHs), aminoglycosides nucleotidyltransferases (ANTs), aminoglycosides acetyltransferases (AACs) that confer resistance to antibiotics such as gentamicin, kanamycin, neomycin, amikacin, and adenylyltransferase (AAD) for streptomycin resistance (53, 73). Also, resistance to aminoglycosides among lactobacilli is generally considered as

intrinsic resistance. The absorption of this type of antibiotic is connected to the transport of electrons mediated by the cytochrome, a system absent in lactobacilli, which therefore determines the inability of the drug to be absorbed by the cell (15, 23, 27, 29, 69). In fact, in several studies, the genes encoding the enzymes for the inactivation of the target antibiotic weren't identified by molecular methods, although the isolates had high MIC values. For example, Guo et al. (2017) (60) showed that out of 14 lactobacilli strains of the *L. casei*, *L. plantarum*, and *L. helveticus* species found to be resistant to kanamycin (MIC> 64 μg/mL), only one strain of *L. helveticus* carried the correspondent *aph* (3″)-III gene (Table 8). None of the streptomycin-resistant strains presented the corresponding determinants *aad*A, *aad*E, *ant* (6). Similarly, Anisimova and Yarullina (2019) (15) did not observe the presence of *aac* (6′)-Ie-*aph* (2″)-Ia, *ant* (6), *aph* (3)-III and *ant* (2″)-I genes in strains resistant to amikacin, kanamycin, and gentamicin. In contrast, the work of de Souza et al. (2019) (68) showed how antibiotic inactivation genes may be present and not always correlated with phenotypic resistance. In the study, two lactobacilli sensitive to gentamicin, kanamycin, and streptomycin, possessed the genes *aac* (6′)-Ii, *ant* (4′)-Ia and *aph* (2)-Ic responsible for resistance to a broad spectrum of aminoglycosides and aph (3)-III for resistance to kanamycin and neomycin (Table 8), demonstrating how the evaluation can be complicated due to the presence of silent or not expressed genes. However, different studies revealed a correspondence between phenotypic and genotypic results, defining that the mechanism of enzymatic inactivation mediated by acquired genes can also occur in lactobacilli. The *aac* (6′)-Ie-*aph* (2)-Ia gene codifies the bifunctional enzyme 6′-*N*-acetyltransferase-2″-O-phosphotransferase, able to confer a broad spectrum of inactivation of all aminoglycosides, apart from streptomycin, and it was found in highly resistant strains of *L. delbrueckii* subsp. *bulgaricus* and *L. plantarum* (14, 58) (Table 8). *aad*A, *aad*E, and *ant* (6) genes, encoding the adenylyltransferase and nucleotidyltransferase enzymes, endow with resistance to streptomycin and was found in strains of *L. paracasei*, *L. casei, L. rhamnsosus*, *L. plantarum,* and *L. delbrueckii* subsp. *bulgaricus,* which showed a resistant phenotype (Table 8).

Table 8 - Aminoglycosides resistance genes found in *Lactobacillus* spp.

Gene	Antibiotic	Phenotype	Species	Reference
aac(6')-Ii	Aminoglycosides	*S*	*L. casei*	(de Souza et al., 2019)
ant(4')-Ia	Aminoglycosides	*S*	*L. casei*	(de Souza et al., 2019)
		S	*L. fermentum*	
aph(2")-Ib	Amminoglicosidi	*R*	*L. plantarum*	(Todorov et al., 2017)
		nd	*L. curvatus*	(de Castilho et al., 2019)
aph(2")-Ic	Aminoglycosides	*nd, R*	*L. curvatus*	(de Castilho et al., 2019; Todorov et al., 2019)
		R	*L. delbrueckii*	(Todorov et al., 2019)
		R	*L. fermentum*	
		S	*L. casei*	(de Souza et al., 2019)
aac(6')-Ie-aph(2")-Ia	Aminoglycosides (Gentamicin)	*R*	*L. delbrueckii* subsp. *bulgaricus*	(Yang e Yu, 2019)
		R	*L. plantarum*	(Todorov et al., 2017; Yang e Yu, 2019)
		nd	*L. curvatus*	(de Castilho et al., 2019)
aph(3')-III	Aminoglycosides (Kanamycin / Neomycin)	*R*	*L. helveticus*	(Guo et al., 2017)
		R	*L. paracasei*	(Ouoba et al., 2008)
		nd	*L. cuvatus*	(de Castilho et al., 2019)
		R, S	*L. casei*	(de Souza et al., 2019; Ouoba et al., 2008)
aadA	Aminoglycosides (Streptomycin)	*R*	*L. paracasei*	(Ouoba et al., 2008)
		R	*L. casei*	
		R	*L. rahmnosus*	(Anisimova and Yarullina, 2019)
aadE	Aminoglycosides (Streptomycin)	*R*	*L. plantarum*	(Anisimova and Yarullina, 2019; Ouoba et al., 2008)
		R	*L. casei*	(Ouoba et al., 2008)
ant(6)	Aminoglycosides (Streptomycin)	*R*	*L. delbrueckii* subsp. *bulgaricus*	(Yang and Yu, 2019)

R: resistant

S: sensitive

ND: not determined

2.3.4 Vancomycin

Lactobacilli are intrinsically resistant to vancomycin, as reported in a study in which 77% of the analysed strains (141/182) presented resistance to vancomycin (65) thanks to their peptidoglycan composition (74). This glycopeptide antibiotic can inhibit the synthesis of the cell wall of Gram-positive bacteria by binding to the D-alanyl-d-alanine precursor of the peptidoglycan, compromising the formation of cross-links (53, 75). Two mechanisms mediate the resistance: by replacing the d-alanine residue at the C-terminus of the peptidoglycan precursor with d-lactate or d-serine, to create a precursor with low affinity for vancomycin, or by preventing synthesis of the D-alanyl-d-alanine bond by eliminating the attack site of the antibiotic (15, 68). Most lactobacilli possess endogenous

enzymes capable of synthesizing d-lactate and binding it to peptidoglycan, thus inducing intrinsic resistance (75).

In *Enterococcus* spp. vancomycin-resistant strains, these mechanisms result from the acquisition of a conjugative plasmid, which includes a cluster containing *van*A, *van*H, *van*R, *van*S, *van*X, *van*Y, and *van*Z genes. Of these, *van*A encodes the enzyme d-alanine-d-lactate ligase, conferring a high level of resistance (75). The vanX gene produces a D-alanyl-d-alanine dipeptidase able to hydrolyze the peptidoglycan dipeptide precursor D-alanyl-d-alanine, eliminating the antibiotic attack site, while *van*H encodes a d-lactate dehydrogenase, which converts pyruvate to d-lactate. The presence of these three genes is necessary to obtain the acquired resistance (15, 36, 75). In support of the intrinsic resistance of lactobacilli to vancomycin, several studies highlighted the presence of resistant isolates lacking the characteristic genes described above. Ouoba et al. (2008) (23) showed that on 16 *Lactobacillus* spp. strains belonging to *L. reuteri*, *Lactiplantibacillus paraplantarum*, *L. plantarum*, *L. fermentum*, *L. salivarius*, *Lactobacillus acidophilus*, *L. rhamnosus*, *L. paracasei,* and *L. casei*, only *L. acidophilus* was sensitive to the antibiotic. The MIC value for the resistant strains exceeded 32 µg/mL but no resistance genes were detected. Similarly Kastner et al. (2006) (27) found that more than 50% of their tested strains were phenotypically resistant, but genetic determinants were absent. Although intrinsic resistance has a minimal chance of horizontal transfer, a case in which one *L. plantarum* strain was able to transfer high phenotypic resistance to *Enterococcus faecalis*, both *in vitro* and *in vivo* conjugation experiments has been reported (21). However, the presence of acquired resistance-related genes is not excluded within lactobacilli (Table 9). Several variants of the gene encoding the enzyme ligase (*van*A, *van*B, *van*C1, *van*C2, *van*C2/C3), which substitutes the terminal residue of d-alanine, were observed in *L. plantarum* (28, 58), *L. curvatus* (29, 56), *L. fermentum* and *L. delbrueckii* (56) and *L. reuteri* (76). The transfer of *van*A from enterococci to a probiotic strain of *L. acidophilus* was highlighted *in vitro* and *in vivo* experiments within the digestive tract of mice (77), highlighting the problem of possible subsequent re-transfer to commensals or pathogens within the human gastrointestinal tract. The *van*X gene, encoding the dipeptidase enzyme, is considered to guarantee lower resistance values than the *van*A and variant genes. It has been found in several *L. plantarum* strains (15, 60, 78).

Table 9 - Vancomycin resistance genes found in *Lactobacillus* spp.

Gene	Coded Enzyme	*Species*	Reference
*van*A	D-alanine-D-lactate ligase	*L. curvatus*	(de Castilho et al., 2019; Todorov et al., 2019)
		L. fermentum	(Todorov et al., 2019)
		L. reuteri	(Dlamini et al., 2019)
		L. plantarum	(Arellano et al., 2019)
		L. garvieae	(de Castilho et al., 2019)
vanB	D-alanine-D-lactate ligase	*L. curvatus*	(de Castilho et al., 2019; Todorov et al., 2019)
		L. plantarum	(Arellano et al., 2019)
vanC1	D-alanine-D-lactate ligase	*L. plantarum*	(Arellano et al., 2019; Todorov et al., 2017)
vanC2	D-alanine-D-lactate ligase	*L. curvatus*	(de Castilho et al., 2019; Todorov et al., 2019)
		L. delbrueckii	(Todorov et al., 2019)
		L. fermentum	
		L. plantarum	(Arellano et al., 2019)
vanC2/C3	D-alanine-D-lactate ligase	*L. plantarum*	(Arellano et al., 2019)
vanX	D-alanyl-D-alanine dipeptidase	*L. brevis*	(Anisimova e Yarullina, 2019)
		L. fermentum	
		L. casei	(Guo et al., 2017)
		L. helveticus	
		L. plantarum	(Anisimova and Yarullina, 2019; Guo et al., 2019; Liu et al., 2009)

2.3.5 B-lactam antibiotics

B-lactam antibiotics are drugs with bactericidal action involving the inhibition of cell wall synthesis. Their structure binds to the transpeptidase enzyme (PBPs - Penicillin Binding Protein), making it no longer available for the formation of cross-links within the peptidoglycan structure. In Gram-positive, resistance is generally mediated by a modification of the antibiotic target molecules by reducing the binding affinity, while the production of β-lactamases resulted in the resistance factor for Gram-negative bacteria (53, 79).

Although lactobacilli are generally considered susceptible to this antibiotics class, such as ampicillin and penicillin (15, 80, 81, 82), some authors reported the presence of ampicillin and cephalosporins resistance in different lactobacilli such as *L. fermentum*, *L. plantarum*, *Levilactobacillus brevis*, *L. salivarius*, *Lactobacillus crispatus* (15, 60, 83, 84) and penicillin G in strains of *L. plantarum* (69). In some cases, resistance is linked to the presence of acquired genes encoding broad spectrum β-lactamase (ESBL - Extended Spectrum β-Lactamase) such as *bla*CTX-M, *bla*SHV, *bla*TEM, and *bla*Z (Table 10) (15, 26, 84). Of these, *bla*CTX-M, which exhibits a high activity spectrum, can be associated with transposons determining its possible spread (84).

Table 10 - Genes of resistance to β-lactam antibiotics found in *Lactobacillus* spp.

Gene	Coded Enzyme	Species	Reference
*bla*CTX-M	ESBL*	*Lactobacillus spp.*	(Khan et al., 2019)
*bla*SHV	ESBL*	*L. brevis ssp. gravesensis*	(Anisimova and Yarullina, 2019)
		L. buchneri	
		L. plantarum	
*bla*TEM	ESBL*	*L. brevis ssp. gravesensis*	(Anisimova and Yarullina, 2019)
		L. brevis	
		L. buchneri	
		L. rhamnosus	
		L. fermentum	
		L. plantarum	
*bla*Z	β-lactamase	*L. plantarum*	(Aquilanti et al., 2007)
*bla*OXA-1	Carbapenemase	*L. brevis ssp. gravesensis*	(Anisimova and Yarullina, 2019)
		L. buchneri	
		L. fermentum	
*bla*OXA-48	Carbapenemase	*L. rhamnosus*	(Hazirolan et al., 2019)

*ESBL= extended spectrum β-lactamase

For the treatment of ESBL producing bacteria, an alternative therapy is the administration of carbapenems (79, 84), although resistance mechanisms also to these antibiotics have been reported, due to the presence of carbapenemase enzymes encoded by genes such as *bla*OXA-48 observed for the first time in *L. rhamnosus* (80), and *bla*OXA-1 in *L. brevis* subsp. *gravesensis*, *Lentilactobacillus buchneri*, and *L. fermentum* (15) (Table 10). Resistance is not always linked to the presence of corresponding genes or β-lactam activity, with uncertainty about the mechanism that leads to the reduction of susceptibility in this microbial genus (83). There is a lack of data in the literature regarding the spread of these determinants from lactobacilli to pathogenic or commensal microorganisms, stressing that further investigation would be desirable.

2.3.6 Ciprofloxacin

About 70% of the tested lactobacilli strains were reported to be resistant to ciprofloxacin, an antibiotic belonging to the fluoroquinolone family whose activity consists in the inhibition of DNA gyrase and DNA topoisomerase IV enzymes, interfering with DNA replication and subsequent microbial growth (14, 53, 60, 69, 85). Moreover, the percentage grows up to 95% when the strains belong to the species *L. plantarum*, *L. reuteri*, *L. salivarius*, *L. brevis*, *L. fermentum,* and *L. rhamnosus* (15, 23, 81).

Within the *Enterococcus* genus, resistance is mediated by the presence of amino acid modifications within the A subunit of the DNA gyrase encoded by the *gyr*A gene, and the C subunit of the topoisomerase IV encoded by the *par*C gene, to reduce the quinolone affinity for the enzymes. The mutations corresponding to this resistance consisted mainly in the substitutions of serine-83 with arginine (*Ser83-to-Arg*), glutamic acid-87 with glycine or lysine (*Glu87-to-Gly* or *Lys*) within the QRDR region (Quinolone Resistance-Determining Region) of the GyrA subunit, and the replacement of serine-80 with leucine or isoleucine (*Ser80-to-Leu* or *Ile*) in the ParC subunit (85). Several authors reported the presence of *gyr*A and *par*C genes within resistant lactobacilli (Table 11), but in none of the cases, the mutations corresponded to the typical ones described above (15, 23), giving evidence of the presence of other amino acid substitutions in the DNA gyrase gene (69; 124). However, it is uncertain whether these substitutions were the cause of the resistance. Consequently, the mechanism that controls resistance to ciprofloxacin in lactobacilli is still unclear.

Table 11 - Ciprofloxacin resistance genes found in *Lactobacillus* spp.

Gene	Species	Reference
*gyr*A	*L. curvatus*	(Hummel et al., 2007)
	L. acidophilus	(Hummel et al., 2007; Ouoba et al., 2008)
	L. reuteri	(Ouoba et al., 2008)
	L. plantarum	(Guo et al., 2017; Hummel et al., 2007; Ouoba et al., 2008)
	L. casei	(Guo et al., 2017)
	L. helveticus	
*par*C	*L. acidophilus*	(Hummel et al., 2007)
	L. curvatus	
	L. buchneri	(Anisimova and Yarullina, 2019)
	L. brevis	
	L. brevis subsp. *gravesensis*	
	L. plantarum	(Anisimova and Yarullina, 2019; Hummel et al., 2007)

2.3.7 Chloramphenicol

Generally, members of the *Lactobacillus* genus are susceptible to the action of chloramphenicol (15, 23, 27, 68), which acts binding to the ribosomal 50 S subunit preventing bacterium protein synthesis (53). However, some studies reported MIC values for this antibiotic above the cut-offs, such as several isolates of *L. plantarum*, with MIC ≥16 μg/mL (28), or a strain of *L. reuteri* for which MIC concentration reached 128 μg/mL (12). In support of the possible resistance to this antibiotic, Campedelli et al. (2019) (65) found resistance to chloramphenicol in 49% of the 182 strains of *Lactobacillus* spp. tested.

Resistance depends on the inactivation of chloramphenicol usually driven by the presence of a chloramphenicol transacetylase enzyme, encoded by genes of the *cat* family (Chloramphenicol Acetyl Transferase) (53). The presence of these genes has been observed in strains of *L. acidophilus* and *L. delbrueckii* subsp. *bulgaricus* isolated from yogurt, *L. plantarum* isolated from salami and green tea, and *L. reuteri* isolated from dog feces (Table 12). However, the phenotypic resistance was not always accompanied by the presence of the *cat* gene, and vice versa. Arellano et al. (2020) (28) lit upon the *cat* gene in two strains of *L. plantarum* susceptible to the antibiotic, an outcome in agreement with the results obtained by Todorov et al. (2017) (58). Similarly, the gene was present, but unexpressed, in one strain of *L. acidophilus* and one of *L. delbrueckii* subsp. *bulgaricus*, resulting in susceptibility to the antibiotic (Table 12), probably due to the presence of mutations in the regulatory region that prevent the expression of the gene (Hummel et al., 2007) (69). The same authors also reported the opposite situation, observing how in *Lactiplantibacillus pentosus* and *L. plantarum* resistant to chloramphenicol, the *cat* gene was not present. A similar result was obtained from the analysis of 43 lactobacilli isolated from dairy products, more than half of which demonstrated resistance to the antibiotic, without however highlighting the related gene (Yang and Yu, 2019) (14). In this case, the mechanism triggering the acquired resistance *cat* is not well understood. As regards the problem of possible horizontal transfer, the presence of the gene on plasmids was observed, consequently, the possible movement cannot be excluded (Egervärn et al., 2009) (12).

Table 12 - Chloramphenicol resistance genes found in *Lactobacillus* spp.

Gene	Species	Phenotype	Reference
cat	*L. acidophilus*	S	(Hummel et al., 2007)
	L. delbrueckii subs. bulgaricus	S	
	C. alimentarius	5/7 R	(Campedelli et al., 2019)
	L. brevis	7/8 R	
	L. buchneri	1/1 S	
	L. casei maniotivorans	1/1 R	
	L. collinoides	1/4 R	
	L. plantarum	3/7 R	
	L. reuteri - vaccinostercus	1/2 R	
	L. sakei	1/2 R	
	L. saivarius	1/3 R	
*cat*A	*L. plantarum*	S	(Todorov et al., 2017)
		S	(Arellano et al., 2019)
cat-TC	*L. reuteri*	R	(Egervarn et al., 2009)

S: susceptible

R: resistant

2.3.8 Other antibiotics - bacitracin, rifampicin, and sulfamethoxazole

Literature counts very few studies analysing resistance and the presence of related resistance genes to antibiotics such as bacitracin, rifampicin, and sulfamethoxazole. In the case of bacitracin, resistance was observed in strains of *L. delbrueckii* subsp. *bulgaricus* and *L. plantarum* isolated from fermented milk (78), in a strain of *L. curvatus* isolated from salami (29), and in strains of *L. plantarum*, *L. pentosus*, *L. paracasei*, *L. rhamnosus*, *L. acidophilus*, *L. sakei,* and *L. curvatus* (86). Some authors, although not evaluating the phenotypic resistance, highlighted in lactobacilli the presence of genes related to antibiotic resistance such as *bcr*B, *bcr*D, and *bcr*R (Table 13). *bcr*B encodes proteins necessary for the drug efflux, *bcr*D encodes an undecaprenol kinase, and *bcr*R identifies a presumed regulatory gene upstream of the *bcr*ABD operon (87). Of the cases reported in Table 13, only Arellano et al. (2020) (28) observed the presence of all three genes of the operon in two strains of *L. plantarum*; in other cases, it would be useful to analyze the phenotypic result to assess whether the presence of a single gene of the operon can affect the susceptibility.

Table 13 - Bacitracin resistance genes found in *Lactobacillus* spp.

Gene	Species	Phenotype	Reference
bcrB	*L. plantarum*	nd	(Arellano et al., 2019)
	L. curvatus	nd	(de Castilho et al., 2019)
	L. casei	nd	(Casarotti et al., 2017; de Souza et al., 2019)
	L. fermentum	nd	(de Souza et al., 2019)
bcrD	*L. plantarum*	nd	(Arellano et al., 2019)
bcrR	*L. plantarum*	nd	(Arellano et al., 2019)
	L. casei	nd	(Casarotti et al., 2017)

nd: not determined

Concerning rifampicin, resistance is generally related to the presence of mutations in the RRDR region (RIF Resistance-Determining Region) of the *rpo*B gene encoding the β subunit of RNA polymerase. In the strains with this mutation, the antibiotic will no longer be able to inhibit the bond between DNA and RNA polymerase, and protein synthesis proceeds (88). Strains of L. *plantarum* and *L. casei* showed resistance to rifampicin (15, 60). In the literature, there is a lack of studies that evaluate the potential presence of mutations within the *rpo*B gene that induces resistance in lactobacilli.

In the case of sulfamethoxazole, resistance within the *Lactobacillus* genus is generally considered intrinsic, probably due to the structure of the cell wall and the impermeability of the membrane (14). This resistance was found in 27 strains of *L. reuteri*, *L. plantarum*, *L. fermentum*, *L. salivarius*, *L. acidophilus*, *L. rhamnosus*, *L. paracasei*, *L. casei,* and *L. delbrueckii* subsp. *bulgaricus* with MIC values even greater than 1024 μg/mL (14, 23).

2.4 Virulence

Lactobacilli are generally considered non-pathogenic for humans. However, some clinical isolates belonging to this heterogenous group have been identified as responsible for health problems and diseases (especially in the case of individuals with previous medical conditions), thanks to the presence of genes capable of encoding virulence factors, which are detailed below.

2.4.1 Adhesion and aggregation factors

The ability to adhere to human tissues and the gastrointestinal tract is the first critical factor for the virulence of various pathogenic microorganisms. Regarding probiotic microorganisms and gut microbiota bacteria, it is usually a positive feature and defines a selection criterion in the evaluation

of new probiotics (25, 76, 89). This property allows the colonization and maintenance of these microorganisms within the gastrointestinal tract. Lactobacilli produce different types of proteins able to adhere to different targets in human tissues and intestinal mucosae such as mucin, collagen, and fibronectin. Several genes coding for these proteins have been reported, as the mucin binding genes *mub* (mucus-binding-protein), *msa* (mannose-specific adhesin) (28,29, 90), and genes encoding proteins able to bind fibrinogen, fibronectin, and collagen such as *fbp*B in *L. acidophilus* (91), *fbp*A in *L. acidophilus* and *L. casei* (90, 92), *cbs*A, *slp*A and *cn*BP in *L. crispatus*, *L. brevis* and *L. reuteri,* respectively (Vélez et al., 2007).

Although the ability to bind fibrinogen is generally considered positive, Collins et al. (2012) (33) reported that *L. salivarius* CCUG_47,825 isolated from a case of septicemia was able, thanks to the presence of a specific protein, to bind fibrinogen and subsequently induce platelet aggregation to a comparable level with *Staphylococcus aureus*. The discovered protein gene was renamed CCUG_2371 (Table 14) and encodes a surface protein rich in serine with similar traits with proteins of *Clostridium perfringens* (38%), *Streptococcus infantarius* (34%), and *Corynebacterium diphtheria* (27%), whose expression is regulated by another gene, CCUG_0873. Collagen binding is also a positive factor in probiotic microorganisms, but a factor associated with pathogenicity in pathogens. The bound is mediated by proteins, such as alpha enolases encoded by *eno* genes (51, 93). These proteins, in pathogenic microorganisms such as *Streptococcus pneumoniae* and *Staphylococcus aureus*, mediate the bound to fibronectin, collagen, and laminin and are also able to bind the plasmin activator (plasminogen), causing uncontrolled proteolysis inducing the possibility of host tissues invasion (51, 94). α-enolases produced by a strain of *L. curvatus* isolated from chicken (*eno* 1, *eno* 2) an a strain of *L. johnsonii* isolated from calf feces (*eno*), shared functional similarity in the traits associated with virulence with α-enolase of *Streptococcus pyogenes*, *Streptococcus pneumoniae,* and *Staphylococcus aureus*, showing the same ability to activate plasminogen and bind laminin, a constituent of various tissues, including heart valves. Further studies are needed to better understand the potential risk connected with these findings (51).

Table 14 - Virulence genes found in *Lactobacillus* spp.

Virulence factor	Gene	Species	Reference
Sex pheromone	*ccf*	*L.curvatus*	(de Castilho et al., 2019)
		L. delbrueckii subsp. bulgaricus	(Casarotti et al., 2017)
		L. casei	
	cob	*L. plantarum*	(Todorov et al., 2017)
		L. delbrueckii subsp. bulgaricus	(Casarotti et al., 2017)
	cpd	*L. plantarum*	(Todorov et al., 2017)
		L. delbrueckii subsp. bulgaricus	(Casarotti et al., 2017)
		L. casei	(Casarotti et al., 2017; de Souza et al., 2019)
Citolisine	*cylA*	*L. curvatus*	(Todorov et al., 2019)
		L. fermentum	(Casarotti et al., 2017)
		L. casei	
		L. mucosae	(de Moraes et al., 2017)
	cylB	*L. paraplantarum*	(Soleymanzadeh et al., 2017)
		L. kefiri	
		L. paracasei	
		L. gasseri	
Adhesion	*ace*	*L. plantarum*	(Soleymanzadeh et al., 2017; Todorov et al., 2017; Todorov et al., 2014)
		L. paracasei	(Soleymanzadeh et al., 2017)
		L. curvatus	(de Castilho et al., 2019)
	CCUG_0873	*L. salivarius 47825*	(Collins et al., 2012)
	CCUG_2731		
	eno	*L. crispatus* ST1	(Antikainen et al., 2007)
		L. johnsonii F133	
Aggregation	*agg*	*L. reuteri*	(Dlamini et al., 2019)
	asa1	*L. plantarum*	(Arellano et al., 2019; Todorov et al., 2017; Todorov et al., 2014)
		L. delbrueckii subsp. bulgaricus	(Casarotti et al., 2017)
		L. curvatus	(de Castilho et al., 2019)
Enterococcal surface protein	*esp*	*L. curvatus*	(de Castilho et al., 2019; Todorov et al., 2019)
		L. casei	(Casarotti et al., 2017)
		L. mucosae	(de Moraes et al., 2017)
		L. delbrueckii	(Todorov et al., 2019)
Endocarditis antigen	*efaA*	*L. kefiri*	(Soleymanzadeh et al., 2017)
		L. paracasei	
		L. gasseri	
		L. plantarum	
		L. mucosae	(de Moraes et al., 2017)
		L. casei	(de Souza et al., 2019)
		L. fermentum	
		L. curvatus	(Todorov et al., 2019)
		L. fermentum	
Gelatinase	*gelE*	*L. gasseri*	(Soleymanzadeh et al., 2017)
		L. plantarum	(Soleymanzadeh et al., 2017; Todorov et al., 2014)
		L. casei	(Casarotti et al., 2017)
		L. mucosae	(de Moraes et al., 2017)
		Lb. delbrueckii	(Casarotti et al., 2017; Todorov et al., 2019)
		L curvatus	(Todorov et al., 2019)
	fsrA	*L. casei*	(Casarotti et al., 2017)
	fsrB	*L. casei*	(Casarotti et al., 2017)
	fsrC	*L. delbrueckii subsp. bulgaricus*	(Casarotti et al., 2017)
		L. casei	
Hyaluronidase	*hyl*	*L. mucosae*	(de Moraes et al., 2017)
		L. plantarum	(Todorov et al., 2014)
		L. delbrueckii subsp. bulgaricus	(Casarotti et al., 2017)
		L. curvatus	(de Castilho et al., 2019)

In some cases, lactobacilli were associated with the presence of virulence genes related to adhesion and colonization properties of *Enterococcus* spp. (Table 14). This may be the result of gene transfer, as these factors are generally encoded by genes located in conjugative plasmids (95, 96). The *ace* gene, implicated in the pathogenesis of *E. faecalis,* was also observed in strains of *L. plantarum*, *L. paracasei,* and *L. curvatus* (Table 14). This gene encodes a surface protein with adhesive

characteristics that mediates adhesion to extracellular matrix proteins, such as type I and IV collagen and laminin, (95, 97, 98). As previously described, the ability to bind these compounds is usually not considered a negative factor for lactobacilli as they are involved in better adhesion and colonization of the gastrointestinal (GI) tract. However, a problem could be due to the presence of enterococcal genes associated with bacterial aggregation such as *agg* and *asa*1. These genes are located on the pAD1 pheromone-inducible plasmid and mediate the production of conjugative aggregates of cells to promote the exchange of genetic material (95, 96, 97). Consequently, the presence of these enterococcal genes in *L. plantarum*, *L. reuteri*, *L. delbrueckii* subsp. *bulgaricus* and *L. curvatus* (Table 14) can be of concern in terms of horizontal transfer of virulence or AR genes to opportunistic pathogens. Other virulence enterococcal genes found in *Lactobacillus* spp. are *efa*A, enterococcal antigen associated with endocarditis, and *esp*, encoding a surface protein, both involved in biofilm formation. *esp* is frequently observed in clinical isolates of *Enterococcus* spp. (99). It is involved in adhesion to human tissues, such as the GI tract and urinary tract, and can mediate the production of biofilm in *Enterococcus* spp. This property is strictly connected to adhesion and resistance to host defence mechanisms, such as phagocytosis, thus promoting pathogenicity (95, 100). It was observed in strains of *L. curvatus*, *L. casei, Limosilactobacillus mucosae,* and *L. delbrueckii* (Table 14). However, the adhesive properties conferred by this gene do not constitute a direct risk in a potential probiotic, but only in the case of transfer to other bacteria. In fact, the ability to adhere to intestinal epithelial cells and to inhibit pathogens' growth in the human gastrointestinal tract (GIT) through competitive exclusion is a desirable feature for probiotic bacteria, as it increases persistence in the GIT and the ability to effectively colonize the intestine (89, 101). Similarly, the presence of the biofilm mediating *efa*A gene, which encodes a specific antigen of enterococci found in clinical isolates of endocarditis cases and urinary tract infections (99, 102), was found in strains of *Lentilactobacillus kefiri*, *L. paracasei*, *Lactobacillus gasseri*, *L. plantarum*, *L. mucosae*, *L. casei* and *Limosilactobacillus fermentum* (Table 14). Being specific of *Enterococcus* spp, the presence of these genes in lactobacilli is certainly the result of horizontal transfer. The problem, therefore, lies in the possibility of re-transfer of these genes to other potentially pathogenic bacteria rather than in the induced capacity to create biofilms.

2.4.2 Sex pheromones

As already mentioned, generally the presence of AR or virulence genes in lactobacilli it is due to a genetic acquisition mechanism mediated by the presence of mobile genetic elements such as plasmids. The ensuing concern is the possibility of retransfer of these determinants to commensal and pathogenic microorganisms. The presence of genes acquired from *Enterococcus* spp. coding for sex

pheromones (*ccf*, *cob*, *cpd*) represents a potential hazard if transferred. Sex pheromones are small peptides produced by Gram-positive microorganisms that work as signal molecules able to mediate the quorum-sensing mechanism (103). Their presence induces conjugation processes mediated by pheromone-inducible plasmids (e.g. pAD1, pPD 1, pCF10), supporting the possibility of transferring determinants that could increase pathogenicity in some bacteria (97, 104). For example, the pAD1 plasmid can carry genes for hemolysin and aggregating substances (96). Sex pheromones are produced by potential recipient cells and received by donor cells, which import the exogenous pheromone, causing the expression of genes involved in the conjugation process of the related pheromone-inducible plasmid. The expression of genes encoding aggregation substances (*agg*, *asa*1) present in plasmids is induced to promote bacterial aggregation by facilitating contact between donor and recipient and, consequently, the passage of the plasmid itself (95, 97, 104). In this way, the strains containing the genes encoding the sex pheromones can acquire the corresponding plasmids and the associated virulence or AR determinants. Several lactobacilli highlighted the presence of the genes *ccf*, *cob,* and *cpd* encoding sex pheromones. They were detected in strains of *L. curvatus*, *L. delbrueckii* subsp. *bulgaricus*, *L. casei,* and *L. plantarum* (Table 14). In some cases, these genes were observed in lactobacilli that also carried the *asa*1 gene responsible for bacterial aggregation activity (58, 89). The presence of these molecules may cause concern given the possibility of inducing the horizontal transfer mechanism by promoting the spread of potentially harmful determinants. Further studies are needed to evaluate the actual role of these molecules within lactobacilli, evaluating the possibility of the latter working as recipients of virulence genes.

2.4.3 Cytolysin toxin

Another factor related to the virulence of pathogenic microorganisms is the ability to produce cytolysin, an exotoxin with β-hemolytic activity. This molecule can be considered an antimicrobial (lantibiotic bacteriocin) given its action against various Gram-positive microorganisms, but also a virulence factor considering the activity against eukaryotic cells such as erythrocytes and human epithelial cells, inducing the onset of infections such as endocarditis (105, 106). This toxin consists of two subunits and its production is delegated to the combined action of eight genes (*cyl*R1, *cyl*R2, *cyl*LL, *cyl*LS, *cyl*M, *cyl*B, *cyl*A, and *cyl*I) transcribed as a single operon. Only if the entire operon is transcribed, the lytic activity is observed. Of the eight genes, *cyl*LS and *cyl*LL are required for the ribosomal synthesis of the small and large subunits, which subsequently undergo a post-translational modification by the *cyl*M gene creating the two products CylLS * and CylLL * with modifications characteristic of lantibiotic bacteriocins. The latter will then be secreted and proteolytically processed by the membrane transporter encoded by the *cyl*B gene creating the two CylLL 'CylLS' subunits,

which will be again subjected to a proteolytic event catalyzed by a serine protease encoded by the *cyl*A gene, generating the two active subunits CylLL ″ CylLS ". The product encoded by the *cyl*I gene allows protection of the producer strain against the lytic activity itself. *cyl*R1 and *cyl*R2 function as repressors, encoding two proteins that repress the transcription of the operon (106). The operon can be located on the same pAD1-sensitive plasmid carrying the *asa*1 adhesive gene (96, 97) or within pathogenicity islands on the associated chromosome to other genes responsible for aggregation and adhesion, such as the *esp* gene (106). Some operon genes described above were observed in lactobacilli. (Table 14). The *cyl*A gene was detected in strains of *L. curvatus*, L. *mucosae*, *L. fermentum,* and *L. casei* isolated from smoked salmon, goat milk, and buffalo mozzarella water, respectively, while *cyl*B gene in *L. paraplantarum*, *L. kefiri*, *L. paracasei,* and *L. gasseri* strains. The reported studies generally evaluated only the presence or absence of the *cyl*A gene without considering the entire operon. However, the activity is linked to several associated genes, as reported by Casarotti et al. (2017) (89) who observed the presence of *cyl*A gene in a strain of *L. casei*, unable to induce hemolysis, as also reported by Todorov et al. (2019) (56) who showed a negative phenotype for a strain of *L. curvatus cyl*A+. Soleymanzadeh et al. (2017) (107) showed that of the four strains found to be *cyl*B + none exhibited hemolytic activity due to the absence of associated genes such as *cyl*A and *cyl*M. In the case reported by de Moraes et al. (2017) (101), the presence of *cyl*A gene, the only gene evaluated, was related to the hemolytic activity for a strain of *L. mucosae.* This does not mean that the mere presence of cylA induced the positive phenotype because other associated genes were not investigated. This suggests that in-depth research should evaluate the presence of all genes present in the operon. However, Casarotti et al. (2017) (89) found the presence of three strains belonging to the species *L. delbrueckii* subsp. *bulgaricus* and a strain of *L. casei* able to induce partial hemolysis but none containing the *cyl*A gene, probably due to other lytic genes.

2.4.4 Gelatinase

Gelatinase is a zinc-dependent extracellular metalloendopeptidase and is considered a virulence factor as it contributes to the degradation of host tissues such as collagen, fibrin, and elastin to supply nutrients to the cell favoring invasion, and it is also involved in biofilm formation (95, 97, 100). It is encoded by the *gel*E gene and generally secreted by clinical isolates belonging to *Enterococcus faecalis* and *Enterococcus faecium,* but it was also found in samples obtained from dairy products (108). It was observed that this peptidase contributed to various disorders such as peritonitis, ulcerative colitis, endocarditis, and irritable bowel disorder in humans and mice (103). The *gel*E gene was discovered in isolates of *L. curvatus* and *L. delbrueckii* from smoked salmon (56), *L. mucosae*, *L. gasseri,* and *L. plantarum* isolated from milk samples (101, 107), *L. casei* and *L. delbrueckii* subsp.

bulgaricus from buffalo mozzarella water (89) and *L. plantarum* isolated from papaya (109) (Table 14). However, the presence of the *gel*E gene alone is not directly correlated to gelatinase activity (56, 101, 107). This is because the expression of this gene is regulated by a quorum-sensing mechanism driven by the *fsr* locus. This mechanism can induce the transcription of specific genes based on the presence and relative concentration in the extracellular space of specific inducing molecules. The *fsr* locus is placed directly in contact with the virulence gene and is made up of three genes: *fsr*A, *fsr*B, *fsr*C. The first works as a regulator for the expression of the *gel*E gene, while *fsr*B encodes a transmembrane transporter able to process and produce the inducing molecule GBAP (Gelatinase Biosynthesis Activating Pheromone), which accumulates in the environment. When the inducer reaches a threshold concentration, the transmembrane histidine protein kinase FsrC, encoded by the *fsr*C gene, perceives the signal induced by GBAP and, after being subjected to phosphorylation, activates the FsrA regulator, which will then subsequently activate the transcription of gelatinase (95, 103). Therefore, the presence of the entire operon is required to obtain a positive GEL + phenotype (108). The same authors underlined how laboratory manipulation, resulting in the loss of the *fsr* operon caused the consecutive loss of gelatinase activity in several enterococcal isolates. Among the published studies, only Casarotti et al. (2017) (89) evaluated the presence of the *gel*E-associated *fsr*A, *fsr*B, and *fsr*C genes in one strain *of L. fermentum*, six strains of *L. casei,* and three strains of *L. delbrueckii* subsp. *bulgaricus* (Table 14), but none of the strains showed the presence of the entire operon.

2.4.5 Hyaluronidase

Hyaluronidase is a degradative enzyme, able to depolymerize the mucopolysaccharides that constitute the human connective tissue, such as hyaluronic acid and cartilage, to provide nourishment to pathogens and promote its diffusion inside the host (95, 97). These enzymes are produced by species belonging to the genera *Staphylococcus* spp, *Streptococcus* spp, *Clostridium* spp. and *Enterococcus* spp. The corresponding *hyl* gene encoding the hyaluronidase enzyme was found mainly in clinical isolates of *E. faecium*, and food isolates of *Enterococcus casseliflavus*, *Enterococcus mundtii,* and *Enterococcus durans* (7, 97). Although there is no evidence in the literature that lactobacilli produces this enzyme, the *hyl* gene was found in strains of *L. mucosae*, *L. plantarum*, *L. delbrueckii* subsp. *bulgaricus* and *L. curvatus* isolated from food (Table 14). The gene presence didn't always result in virulence induction, so the expression evaluation is fundamental. However, the major concern is the possible re-transfer of the gene to pathogenic microorganisms since, in clinical isolates of *E. faecium*, the *hyl* gene is positioned on a conjugative plasmid (7).

2.4.6 Glycosidase and arylamidase

As previously reported, lactobacilli were responsible for infectious endocarditis (30, 37, 39, 42, 46). This disorder involves bacterial infection of the endocardial surface through initial bacteremia induced by bacteria that can subsequently adhere to and colonize the tissue of the heart valve (fibrinogen, laminin, collagen). Consequently, the formation of the so-called vegetations, infected masses formed by the deposition of successive cycles of platelet aggregations, which create a protective matrix for the microorganisms capable of proliferating, occurs (110).

It was observed that some species of lactobacilli (*L. rhamnosus*, *L. paracasei*, *L. casei*, *L. salivarius*) had factors favoring the potential colonization of cardiac tissue by adhering to collagen, fibronectin, and fibrinogen, and mediating platelet aggregation (33, 48, 52, 111). Colonization depends also on other factors, such as the ability to obtain nutrients and to evade the host's defence mechanisms. In this regard, Oakey et al. (1995) observed the ability to produce characteristic enzymes, as glycosidase and arylamidase in lactobacilli isolates from cases of endocarditis (Table 15). They highlighted that the combination of *N*-acetyl-β-d-glucosaminidase and α-d-galactosidase production was a characteristic feature of clinical isolates. Their action is delegated to the release of carbohydrates from glycoproteins allowing the inflow of nutrients to the bacterium. They also observed the presence of arylamidase with fibrinolytic and pro-and anticoagulant activity, with its main activities in the evasion of host defense mechanisms. Among these, the presence of (Ca)-like C protein, a human physiological anticoagulant with fibrinolytic activity able to convert plasminogen into plasmin causing the invasion of vegetation (112) was found. The same activity was obtained by the presence of a protein, produced by clinically isolated strains of *L. rhamnosus* and *L. paracasei,* similar to the proteolytic enzyme of plasma kallikrein origin, which is also able to activate the Hageman factor (XII), the zymogenic form of a serine protease with action in the early stages of coagulation (113), thus inducing pro-coagulant activity. This ability determines the expansion of the vegetation around the bacterium giving more resistance from the host's defense mechanisms. The pro-coagulant activity has also been associated with the presence of an Xa-like activated factor, a plasma glycoprotein involved in blood clotting responsible for the conversion of prothrombin into thrombin, which subsequently catalyzes the conversion of fibrinogen into fibrin, a component of the clot (114).

Table 15 - Enzymes found in *Lactobacillus* spp. associated with infectious endocarditis

Enzyme		Role	Species	Reference
Glycosidase	*N-acetyl-β-D-glucosaminidase*	Release carbohydrates from glycoproteins favoring the inflow of nutrients to the bacterium, an important factor for the microbial colonization of thrombotic vegetation	*L. rhamnosus* *L. paracasei* *L. acidophilus*	(Oakey et al., 1995)
	α-*D-galattosidase*			
Protease (Arylamidase)	*Activated factor X*	Pro-coagulant activity		
	(Xa)-like	Aids evasion from the host's defense mechanism		
	Activated protein C	Fibrinolithic activity		
	(Ca)-like	Aids evasion from the host's defense mechanism		
	Hageman factor- like	Pro-coagulant activity		
	Kallikrein-like	Fibrinolithic and indirect ro-coagulant activity		

2.5 Conclusions

Antibiotic resistence and virulence factors are important aspects to be considerend in lactic acid bacteria due to their important role in food production. After a critical reading, it emerged that several studies highlighted the lack of standards in the field of susceptibility testing and the relative definition of the cut-offs for different species of lactobacilli (14, 15, 26, 69, 115, 116). As regards the tests for the assessment of susceptibility, for example, possible interferences in the determination of the MIC value related to the medium used were underlined (117). Klare et al. (2005) (118) developed and tested a specific medium (LSM) for lactic acid bacteria (LAB) able to provide optimal results in terms of growth support and a correct indication of the MIC value. However, the latter has not been included in a standardized method, consequently, in many studies, the De Man, Rogosa, and Sharpe (MRS) medium is still used, of which components can cause potential interference with specific antibiotics such as trimethoprim and sulfonamides (118). Furthermore, the low pH of the MRS medium (pH 6.2 ± 0.2) may be responsible for the reduction of the activity of some antibiotics such as aminoglycosides (optimal pH 7.8), while the pH of the LMS medium, adjusted to pH 6.8, appears not to interfere (118, 119). A further problem in the sensitivity analysis is the determination of the MIC cut-off values of different lactobacilli, which is important since it can affect the decision to consider a bacterium susceptible or resistant. Hummel et al. (2007) underline this problem by presenting discordant results regarding the antibiotic resistance of various lactic bacteria considering the cut-offs dictated by FEEDAP (120), and European Commission (121), and Danielsen and Wind (2003) (86). Lactobacilli are generally considered to be more resistant to aminoglycosides and vancomycin and susceptible to erythromycin, β-lactam antibiotics, chloramphenicol, and tetracycline. However,

strains resistant to the latter antibiotics have also been identified with the presence of the related genes. Resistance to tetracycline, erythromycin, and aminoglycosides are the most discussed, and the *tet*(M) and *erm*(B) genes the most observed, studied and evaluated also considering their possible transfer. From the analyzed data, it emerged that *L. plantarum* is one of the most documented species characterized by different resistance traits and by the presence of the corresponding potentially transmissible genes. It is also important to remember the possible presence, within the same strain, of resistance to multiple antibiotics, as highlighted by Campedelli et al. (2019) (65) who noted multiple resistance in 152 of the 182 strains analyzed (84%). Finally, for a complete evaluation of the safety of lactobacilli, it would be necessary to take into consideration all the factors reported above, both virulence traits and AR.

Considering AR, it is important to investigate on all the possible determinants to avoid confusing an extrinsic resistance, with the relative concern of the possibility of transfer, with an intrinsic resistance, much less problematic.

Regarding virulence factors, although the known genetic determinants of *lactobacilli* are limited, in this review it is exposed that there is the possibility of acquiring virulence genes from pathogenic microorganisms, with the possibility of consequent clinical problems, in particular in immunocompromised patients.

As reported by EFSA, whole genome sequencing can be a valid technique to screen for bacteria intentionally used in the food chain (2) but to date, there is no specific genetic database for lactobacilli associated with a safety assessment, and existing databases tend to focus primarily on pathogens. For example, in some studies, the VFDB database (Virulence Factors of Pathogenic Bacteria) (122) built to analyze virulence factors in genomes of 32 well-established human pathogen genera of which genomes are not part of *Lactobacillus* spp, is mistakenly used. This leads to misleading results, as factors that can induce virulence in certain pathogens are not necessarily given the same result about lactobacilli. For example, Zhang et al. (2012) (123) conducted the safety assessment of L. plantarum JDM1 using this database, resulting in the presence of 126 virulence genes, subsequently not considered problematic as they do not encode toxins or proteins of invasion. This underlines the usefulness of the presence of potential virulence and AR-related gene lists for lactobacilli.

2.6 References

1. EFSA, 2007. Introduction of a Qualified Presumption of Safety (QPS) approach for assessment of selected microorganisms referred to EFSA - Opinion of the Scientific Committee. EFSA J. 5, 1–16. https://doi.org/10.2903/j.efsa.2007.587
2. EFSA, 2021. EFSA statement on the requirements for whole genome sequence analysis of microorganisms intentionally used in the food chain. EFSA J. 19. https://doi.org/10.2903/j.efsa.2021.6506
3. Imperial, I.C.V.J., Ibana, J.A., 2016. Addressing the antibiotic resistance problem with probiotics: Reducing the risk of its double-edged sword effect. Front. Microbiol. 7, 1–10. https://doi.org/10.3389/fmicb.2016.01983
4. Drago, L., Rodighiero, V., Mattina, R., Toscano, M., De Vecchi, E., 2011. In vitro selection and transferability of antibiotic resistance in the probiotic strain lactobacillus reuteri DSM 17938. J. Chemother. 23, 371–373. https://doi.org/10.1179/joc.2011.23.6.371
5. Zheng, J., Wittouck, S., Salvetti, E., Franz, C.M.A.P., Harris, H.M.B., Mattarelli, P., O'toole, P.W., Pot, B., Vandamme, P., Walter, J., Watanabe, K., Wuyts, S., Felis, G.E., Gänzle, M.G., Lebeer, S., 2020. 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. 70, 2782–2858. https://doi.org/10.1099/ijsem.0.004107
6. Lerner, A., Shoenfeld, Y., Matthias, T., 2019. Probiotics: If it does not help it does not do any harm. really? Microorganisms 7, 1–20. https://doi.org/10.3390/microorganisms7040104
7. Van Reenen, C.A., Dicks, L.M.T., 2011. Horizontal gene transfer amongst probiotic lactic acid bacteria and other intestinal microbiota: What are the possibilities? A review. Arch. Microbiol. 193, 157–168. https://doi.org/10.1007/s00203-010-0668-3
8. Ma, Q., Fu, Y., Sun, H., Huang, Y., Li, L., Yu, Q., Dinnyes, A., Sun, Q., 2017. Antimicrobial resistance of Lactobacillus spp. from fermented foods and human gut. Lwt 86, 201. https://doi.org/10.1016/j.lwt.2017.07.059
9. Willis, C., 2000. Antibiotics in the food chain: Their impact on the consumer. Rev. Med. Microbiol. https://doi.org/10.1097/00013542-200011030-00005
10. Founou, L.L., Founou, R.C., Essack, S.Y., 2016. Antibiotic resistance in the food chain: A developing country-perspective. Front. Microbiol. 7, 1–19. https://doi.org/10.3389/fmicb.2016.01881
11. Wang, H.H., Manuzon, M., Lehman, M., Wan, K., Luo, H., Wittum, T.E., Yousef, A., Bakaletz, L.O., 2006. Food commensal microbes as a potentially important avenue

intransmitting antibiotic resistance genes. FEMS Microbiol. Lett. 254, 226–231. https://doi.org/10.1111/j.1574-6968.2005.00030.x

12. Egervärn, M., Roos, S., Lindmark, H., 2009. Identification and characterization of antibiotic resistance genes in Lactobacillus reuteri and Lactobacillus plantarum. J. Appl. Microbiol. 107, 1658–1668. https://doi.org/10.1111/j.1365-2672.2009.04352.x
13. Gevers, D., Danielsen, M., Huys, G., Swings, J., 2003a. Molecular characterization of tet(M) genes in Lactobacillus isolates from different types of fermented dry sausage. Appl. Environ. Microbiol. 69, 1270–1275. https://doi.org/10.1128/AEM.69.2.1270-1275.2003
14. Yang, C., Yu, T., 2019. Characterization and transfer of antimicrobial resistance in lactic acid bacteria from fermented dairy products in China. J. Infect. Dev. Ctries. 13, 137–148. https://doi.org/10.3855/jidc.10765
15. Anisimova, E.A., Yarullina, D.R., 2019. Antibiotic Resistance of LACTOBACILLUS Strains. Curr. Microbiol. 76, 1407–1416. https://doi.org/10.1007/s00284-019-01769-7
16. Hammad, A.M., Shimamoto, T., 2010. Towards a compatible probiotic-antibiotic combination therapy: Assessment of antimicrobial resistance in the Japanese probiotics. J. Appl. Microbiol. 109, 1349–1360. https://doi.org/10.1111/j.1365-2672.2010.04762.x
17. EFSA, 2012. Guidance on the assessment of bacterial susceptibility to antimicrobials of human and veterinary importance. EFSA J. 10, 1–10. https://doi.org/10.2903/j.efsa.2012.2740
18. Lerner, A., Matthias, T., Aminov, R., 2017. Potential effects of horizontal gene exchange in the human gut. Front. Immunol. 8, 1–14. https://doi.org/10.3389/fimmu.2017.01630
19. Broaders, E., Gahan, C.G.M., Marchesi, J.R., 2013. Mobile genetic elements of the human gastrointestinal tract: Potential for spread of antibiotic resistance genes. Gut Microbes 4, 271–280. https://doi.org/10.4161/gmic.24627
20. Gevers, D., Huys, G., Swings, J., 2003b. In vitro conjugal transfer of tetracycline resistance from Lactobacillus isolates to other Gram-positive bacteria. FEMS Microbiol. Lett. 225, 125–130. https://doi.org/10.1016/S0378-1097(03)00505-6
21. Preethi, C., Thumu, S.C.R., Halami, P.M., 2017. Occurrence and distribution of multiple antibiotic-resistant Enterococcus and Lactobacillus spp. from Indian poultry: in vivo transferability of their erythromycin, tetracycline and vancomycin resistance. Ann. Microbiol. 67, 395–404. https://doi.org/10.1007/s13213-017-1270-6
22. Thumu, S.C.R., Halami, P.M., 2019. Conjugal transfer of erm(B) and multiple tet genes from Lactobacillus spp. to bacterial pathogens in animal gut, in vitro and during food fermentation. Food Res. Int. 116, 1066–1075. https://doi.org/10.1016/j.foodres.2018.09.046

23. Ouoba, L.I.I., Lei, V., Jensen, L.B., 2008. Resistance of potential probiotic lactic acid bacteria and bifidobacteria of African and European origin to antimicrobials: Determination and transferability of the resistance genes to other bacteria. Int. J. Food Microbiol. 121, 217–224. https://doi.org/10.1016/j.ijfoodmicro.2007.11.018
24. Zonenschain, D., Rebecchi, A., Morelli, L., 2009. Erythromycin- and tetracycline-resistant lactobacilli in Italian fermented dry sausages. J. Appl. Microbiol. 107, 1559–1568. https://doi.org/10.1111/j.1365-2672.2009.04338.x
25. Marosevic, D., Kaevska, M., Jaglic, Z., 2017. Resistance to the tetracyclines and macrolide-lincosamide-streptogramin group of antibiotics and its genetic linkage – A review. Ann. Agric. Environ. Med. 24, 338–344. https://doi.org/10.26444/aaem/74718
26. Aquilanti, L., Garofalo, C., Osimani, A., Silvestri, G., Vignaroli, C., Clementi, F., 2007. Isolation and molecular characterization of antibiotic-resistant lactic acid bacteria from poultry and swine meat products. J. Food Prot. 70, 557–565. https://doi.org/10.4315/0362-028X-70.3.557
27. Kastner, S., Perreten, V., Bleuler, H., Hugenschmidt, G., Lacroix, C., Meile, L., 2006. Antibiotic susceptibility patterns and resistance genes of starter cultures and probiotic bacteria used in food. Syst. Appl. Microbiol. 29, 145–155. https://doi.org/10.1016/j.syapm.2005.07.009
28. Arellano, K., Vazquez, J., Park, H., Lim, J., Ji, Y., Kang, H.J., Cho, D., Jeong, H.W., Holzapfel, W.H., 2020. Safety Evaluation and Whole-Genome Annotation of Lactobacillus plantarum Strains from Different Sources with Special Focus on Isolates from Green Tea. Probiotics Antimicrob. Proteins 12, 1057–1070. https://doi.org/10.1007/s12602-019-09620-y
29. de Castilho, N.P.A., Nero, L.A., Todorov, S.D., 2019. Molecular screening of beneficial and safety determinants from bacteriocinogenic lactic acid bacteria isolated from Brazilian artisanal calabresa. Lett. Appl. Microbiol. 69, 204–211. https://doi.org/10.1111/lam.13194
30. Aaron, J.G., Sobieszczyk, M.E., Weiner, S.D., Whittier, S., Lowy, F.D., 2017. Lactobacillus rhamnosus endocarditis after upper endoscopy. Open Forum Infect. Dis. 4, 4–6. https://doi.org/10.1093/ofid/ofx085
31. Campagne, J., Guichard, J.F., Moulhade, M.C., Kawski, H., Maurier, F., 2020. Lactobacillus endocarditis: a case report in France and literature review. IDCases 21, e00811. https://doi.org/10.1016/j.idcr.2020.e00811
32. Chery, J., Dvoskin, D., Morato, F.P., Fahoum, B., 2013. Lactobacillus fermentum, a pathogen in documented cholecystitis. Int. J. Surg. Case Rep. 4, 662–664. https://doi.org/10.1016/j.ijscr.2013.04.034

33. Collins, J., van Pijkeren, J.P., Svensson, L., Claesson, M.J., Sturme, M., Li, Y., Cooney, J.C., van Sinderen, D., Walker, A.W., Parkhill, J., Shannon, O., O'Toole, P.W., 2012. Fibrinogen-binding and platelet-aggregation activities of a Lactobacillus salivarius septicaemia isolate are mediated by a novel fibrinogen-binding protein. Mol. Microbiol. 85, 862–877. https://doi.org/10.1111/j.1365-2958.2012.08148.x
34. Darbro, B.W., Petroelje, B.K., Doern, G. V., 2009. Lactobacillus delbrueckii as the cause of urinary tract infection. J. Clin. Microbiol. 47, 275–277. https://doi.org/10.1128/JCM.01630-08
35. Groga-Bada, P., Mueller, I.I., Foschi, F., Gawaz, M., Eick, C., 2018. Mitral Valve Endocarditis due to Lactobacillus. Case Rep. Med. 2018, 38–40. https://doi.org/10.1155/2018/8613948
36. Lee, C., Wongjarupong, N., Narayan, M., Melzer, A.C., 2020. Acute Cholecystitis Complicated with Gallbladder Perforation and Peritonitis Caused by Vancomycin-resistant Lactobacillus paracasei. Cureus 12. https://doi.org/10.7759/cureus.7476
37. Martinez, R.M., Hulten, K.G., Bui, U., Clarridge, J.E., 2014. Molecular analysis and clinical significance of lactobacillus spp. Recovered from clinical specimens presumptively associated with disease. J. Clin. Microbiol. 52, 30–36. https://doi.org/10.1128/JCM.02072-13
38. Naqvi, S.S.B., Nagendra, V., Hofmeyr, A., 2018. Probiotic related Lactobacillus rhamnosus endocarditis in a patient with liver cirrhosis. IDCases 13, e00439. https://doi.org/10.1016/j.idcr.2018.e00439
39. Patnaik, S., Davila, C.D., Chennupati, A., Rubin, A., 2015. Endocarditis of the native aortic valve caused by Lactobacillus jensenii. BMJ Case Rep. 2015, 3–6. https://doi.org/10.1136/bcr-2014-206288
40. Recio, R., Chaves, F., Reyes, C.A., Orellana, M.Á., 2017. Infective endocarditis due to Lactobacillus rhamnosus : Risks of probiotic consumption in a patient with structural heart disease. Enfermedades Infecc. y Microbiol. Clin. (English ed.) 35, 609–610. https://doi.org/10.1016/j.eimce.2017.10.008
41. Rossi, F., Amadoro, C., Colavita, G., 2019. Members of the lactobacillus genus complex (LGC) as opportunistic pathogens: A review. Microorganisms 7. https://doi.org/10.3390/microorganisms7050126
42. Stroupe, C., Pendley, J., Isang, E., Helms, B., 2017. Persistent bacteremia secondary to delayed identification of Lactobacillus in the setting of mitral valve endocarditis. IDCases 10, 132–134. https://doi.org/10.1016/j.idcr.2017.10.002

43. Tavernese, A., Stelitano, M., Mauceri, A., Mollace, R., Uccello, G., Romeo, F., Cammalleri, V., 2020. Progression of Lactobacillus plantarum prosthetic valve endocarditis followed by transesophageal echocardiogram. Int. J. Infect. Dis. 97, 160–161. https://doi.org/10.1016/j.ijid.2020.05.067
44. Tena, D., Martínez, N.M., Losa, C., Fernández, C., Medina, M.J., Sáez-Nieto, J.A., 2013. Acute acalculous cholecystitis complicated with peritonitis caused by Lactobacillus plantarum. Diagn. Microbiol. Infect. Dis. 76, 510–512. https://doi.org/10.1016/j.diagmicrobio.2013.03.018
45. Wallet, F., Dessein, R., Armand, S., Courcol, R.J., 2002. Molecular diagnosis of endocarditis due to Lactobacillus casei subsp. rhamnosus. Clin. Infect. Dis. 35, 117–119. https://doi.org/10.1086/344181
46. Zeba, F., Yirerong, J., Assali, M., Tewary, G., Noska, A., 2018. A Double Whammy: Lactobacillus acidophilus Bacteremia and Subsequent Lactobacillus rhamnosus Prosthetic Valve Infective Endocarditis in an Elderly Diabetic Patient. R. I. Med. J. (2013) 101, 32–35.
47. Cannon, J.P., Lee, T.A., Bolanos, J.T., Danziger, L.H., 2005. Pathogenic relevance of Lactobacillus: A retrospective review of over 200 cases. Eur. J. Clin. Microbiol. Infect. Dis. 24, 31–40. https://doi.org/10.1007/s10096-004-1253-y
48. Harty, D.W.S., Oakey, H.J., Patrikakis, M., Hume, E.B.H., Knox, K.W., 1994. Pathogenic potential of lactobacilli. Int. J. Food Microbiol. 24, 179–189. https://doi.org/10.1016/0168-1605(94)90117-1
49. Yelin, I., Flett, K.B., Merakou, C., Mehrotra, P., Stam, J., Snesrud, E., Hinkle, M., Lesho, E., McGann, P., McAdam, A.J., Sandora, T.J., Kishony, R., Priebe, G.P., 2019. Genomic and epidemiological evidence of bacterial transmission from probiotic capsule to blood in ICU patients. Nat. Med. 25, 1728–1732. https://doi.org/10.1038/s41591-019-0626-9
50. Oakey, H.J., Harty, D.W.S., Knox, K.W., 1995. Enzyme production by lactobacilli and the potential link with infective endocarditis. J. Appl. Bacteriol. 78, 142–148. https://doi.org/10.1111/j.1365-2672.1995.tb02834.x
51. Antikainen, J., Kuparinen, V., Lähteenmäki, K., Korhonen, T.K., 2007. Enolases from Gram-positive bacterial pathogens and commensal lactobacilli share functional similarity in virulence-associated traits. FEMS Immunol. Med. Microbiol. 51, 526–534. https://doi.org/10.1111/j.1574-695X.2007.00330.x
52. Harty, D.W.S., Patrikakis, M., Hume, E.B.H., Oakey, H.J., Knox, K.W., 1993. The aggregation of human platelets by Lactobacillus species. J. Gen. Microbiol. 139, 2945–2951. https://doi.org/10.1099/00221287-139-12-2945

53. Kapoor, G., Saigal, S., Elongavan, A., 2017. Action and resistance mechanisms of antibiotics: A guide for clinicians. J. Anaesthesiol. Clin. Pharmacol. 33, 46–50. https://doi.org/10.4103/joacp.JOACP
54. Roberts, M.C., Schwarz, S., 2009. Tetracycline and Chloramphenicol Resistance Mechanisms. Antimicrob. Drug Resist. 183–193. https://doi.org/10.1007/978-1-59745-180-2_15
55. Chang, Y.C., Tsai, C.Y., Lin, C.F., Wang, Y.C., Wang, I.K., Chung, T.C., 2011. Characterization of tetracycline resistance lactobacilli isolated from swine intestines at western area of Taiwan. Anaerobe 17, 239–245. https://doi.org/10.1016/j.anaerobe.2011.08.001
56. Todorov, S.D., Tome, E., Nero, L.A., 2019. Not everything is a question of reputation: Safety of bacteriocinogenic lab isolated from smoked salmon. Chem. Eng. Trans. 75, 451–456. https://doi.org/10.3303/CET1975076
57. Devirgiliis, C., Coppola, D., Barile, S., Colonna, B., Perozzi, G., 2009. Characterization of the Tn916 conjugative transposon in a food-borne strain of Lactobacillus paracasei. Appl. Environ. Microbiol. 75, 3866–3871. https://doi.org/10.1128/AEM.00589-09
58. Todorov, S.D., Perin, L.M., Carneiro, B.M., Rahal, P., Holzapfel, W., Nero, L.A., 2017. Safety of Lactobacillus plantarum ST8Sh and Its Bacteriocin. Probiotics Antimicrob. Proteins 9, 334–344. https://doi.org/10.1007/s12602-017-9260-3
59. Gfeller, K.Y., Roth, M., Meile, L., Teuber, M., 2003. Sequence and genetic organization of the 19.3-kb erythromycin- and dalfopristin-resistance plasmid pLME300 from Lactobacillus fermentum ROT1. Plasmid 50, 190–201. https://doi.org/10.1016/j.plasmid.2003.08.001
60. Guo, H., Pan, L., Li, L., Lu, J., Kwok, L., Menghe, B., Zhang, H., Zhang, W., 2017. Characterization of Antibiotic Resistance Genes from Lactobacillus Isolated from Traditional Dairy Products. J. Food Sci. 82, 724–730. https://doi.org/10.1111/1750-3841.13645
61. Ammor, M.S., Gueimonde, M., Danielsen, M., Zagorec, M., Van Hoek, A.H.A.M., De Los Reyes-Gavilán, C.G., Mayo, B., Margolles, A., 2008. Two different tetracycline resistance mechanisms, plasmid-carried tet(L) and chromosomally located transposon-associated tet(M), coexist in Lactobacillus sakei rits 9. Appl. Environ. Microbiol. 74, 1394–1401. https://doi.org/10.1128/AEM.01463-07
62. Zhang, Y., Zhang, Q., 2019. Relationship between tetracycline antibiotic susceptibility and genotype in oral cavity Lactobacilli clinical isolates. Antimicrob. Resist. Infect. Control 8, 1–8. https://doi.org/10.1186/s13756-019-0483-9

63. Ammor, M.S., Flórez, A.B., Van Hoek, A.H.A.M., De Los Reyes-Gavilán, C.G., Aarts, H.J.M., Margolles, A., Mayo, B., 2007. Molecular characterization of intrinsic and acquired antibiotic resistance in lactic acid bacteria and bifidobacteria. J. Mol. Microbiol. Biotechnol. 14, 6–15. https://doi.org/10.1159/000106077
64. Van Hoek, A.H.A.M., Mayrhofer, S., Domig, K.J., Flórez, A.B., Ammor, M.S., Mayo, B., Aarts, H.J.M., 2008. Mosaic tetracycline resistance genes and their flanking regions in Bifidobacterium thermophilum and Lactobacillus johnsonii. Antimicrob. Agents Chemother. 52, 248–252. https://doi.org/10.1128/AAC.00714-07
65. Campedelli, I., Mathur, H., Salvetti, E., Rea, M.C., Torriani, S., Ross, R.P., Hill, C., O'Toole, P.W., 2019. Genus-Wide Assessment of Antibiotic Resistance in Lactobacillus spp. Appl. Enveronmental Microbiol. 85, 1–21. https://doi.org/https://doi .org/10.1128/AEM.01738-18
66. Cauwerts, K., Pasmans, F., Devriese, L.A., Haesebrouck, F., Decostere, A., 2006a. Cloacal Lactobacillus isolates from broilers often display resistance toward tetracycline antibiotics. Microb. Drug Resist. 12, 284–288. https://doi.org/10.1089/mdr.2006.12.284
67. Leclercq, R., 2002. Mechanisms of resistance to macrolides and lincosamides: Nature of the resistance elements and their clinical implications. Clin. Infect. Dis. 34, 482–492. https://doi.org/10.1086/324626
68. de Souza, B.M.S., Borgonovi, T.F., Casarotti, S.N., Todorov, S.D., Penna, A.L.B., 2019. Lactobacillus casei and Lactobacillus fermentum Strains Isolated from Mozzarella Cheese: Probiotic Potential, Safety, Acidifying Kinetic Parameters and Viability under Gastrointestinal Tract Conditions. Probiotics Antimicrob. Proteins 11, 382–396. https://doi.org/10.1007/s12602-018-9406-y
69. Hummel, A.S., Hertel, C., Holzapfel, W.H., Franz, C.M.A.P., 2007. Antibiotic resistances of starter and probiotic strains of lactic acid bacteria. Appl. Environ. Microbiol. 73, 730–739. https://doi.org/10.1128/AEM.02105-06
70. Feld, L., Bielak, E., Hammer, K., Wilcks, A., 2009. Characterization of a small erythromycin resistance plasmid pLFE1 from the food-isolate Lactobacillus plantarum M345. Plasmid 61, 159–170. https://doi.org/10.1016/j.plasmid.2009.01.002
71. Egervärn, M., Lindmark, H., Olsson, J., Roos, S., 2010. Transferability of a tetracycline resistance gene from probiotic Lactobacillus reuteri to bacteria in the gastrointestinal tract of humans. Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol. 97, 189–200. https://doi.org/10.1007/s10482-009-9401-0
72. Cauwerts, K., Pasmans, F., Devriese, L.A., Martel, A., Haesebrouck, F., Decostere, A., 2006b. Cloacal Lactobacillus isolates from broilers show high prevalence of resistance towards

macrolide and lincosamide antibiotics. Avian Pathol. 35, 160–164. https://doi.org/10.1080/03079450600598137

73. Magalhães, M.L., Blanchard, J.S., 2019. Aminoglycosides: Mechanisms of Action and Resistance. Amin. Antibiot. A Guid. To Ther. 23–44. https://doi.org/10.1201/9780429280184-2
74. Zhang, S., Oh, J.-H., Alexander, L.M., Özçam, M., Van Pijkeren, J.P., 2018. d-Alanyl-d-Alanine Ligase as a Broad-Host-Range Counterselection Marker in Vancomycin-Resistant Lactic Acid Bacteria 200, 1–15. https://doi.org/https://doi.org/10.1128/JB.00607-17
75. Deghorain, M., Goffin, P., Fontaine, L., Mainardi, J.L., Daniel, R., Errington, J., Hallet, B., Hols, P., 2007. Selectivity for D-lactate incorporation into the peptidoglycan precursors of Lactobacillus plantarum: Role of Aad, a VanX-like D-alanyl-D-alanine dipeptidase. J. Bacteriol. 189, 4332–4337. https://doi.org/10.1128/JB.01829-06
76. Dlamini, Z.C., Langa, R.L.S., Aiyegoro, O.A., Okoh, A.I., 2019. Safety Evaluation and Colonisation Abilities of Four Lactic Acid Bacteria as Future Probiotics. Probiotics Antimicrob. Proteins 11, 397–402. https://doi.org/10.1007/s12602-018-9430-y
77. Mater, D.D.G., Langella, P., Corthier, G., Flores, M.J., 2007. A probiotic Lactobacillus strain can acquire vancomycin resistance during digestive transit in mice. J. Mol. Microbiol. Biotechnol. 14, 123–127. https://doi.org/10.1159/000106091
78. Liu, C., Zhang, Z.Y., Dong, K., Yuan, J.P., Guo, X.K., 2009. Antibiotic resistance of probiotic strains of lactic acid bacteria isolated from marketed foods and drugs. Biomed. Environ. Sci. 22, 401–412. https://doi.org/10.1016/S0895-3988(10)60018-9
79. Naas, T., Cuzon, G., Bogaerts, P., Glupczynski, Y., Nordmann, P., 2011. Evaluation of a DNA microarray (check-MDR CT102) for rapid detection of TEM, SHV, and CTX-M extended-spectrum β-lactamases and of KPC, OXA-48, VIM, IMP, and NDM-1 carbapenemases. J. Clin. Microbiol. 49, 1608–1613. https://doi.org/10.1128/JCM.02607-10
80. Hazlrolan, G., Gündoğdu, A., Nigiz, S., Altun, B., Gür, D., 2019. Presence of OXA-48 Gene in a Clinical Isolate of Lactobacillus rhamnosus. Foodborne Pathog. Dis. 16, 840–843. https://doi.org/10.1089/fpd.2019.2651
81. Klein, G., 2011. Antibiotic resistance and molecular characterization of probiotic and clinical Lactobacillus strains in relation to safety aspects of probiotics. Foodborne Pathog. Dis. 8, 267–281. https://doi.org/10.1089/fpd.2010.0672
82. Saini, K., Tomar, S.K., 2017. In vitro evaluation of probiotic potential of Lactobacillus cultures of human origin capable of selenium bioaccumulation. LWT - Food Sci. Technol. 84, 497–504. https://doi.org/10.1016/j.lwt.2017.05.034

83. Dec, M., Nowaczek, A., Stępień-Pyśniak, D., Wawrzykowski, J., Urban-Chmiel, R., 2018. Identification and antibiotic susceptibility of lactobacilli isolated from turkeys. BMC Microbiol. 18, 1–14. https://doi.org/10.1186/s12866-018-1269-6
84. Khan, U., Afsana, S., Kibtia, M., Hossain, M., Choudhury, N., Ahsan, C.R., 2019. Presence of blactx-m antibiotic resistance gene in lactobacillus spp. Isolated from hirschsprung diseased infants with stoma. J. Infect. Dev. Ctries. 13, 426–433. https://doi.org/10.3855/jidc.10968
85. Petersen, A., Jensen, L.B., 2004. Analysis of gyrA and parC mutations in enterococci from environmental samples with reduced susceptibility to ciprofloxacin. FEMS Microbiol. Lett. 231, 73–76 https://doi.org/10.1016/S0378-1097(03)00929-7
86. Danielsen, M., Wind, A., 2003. Susceptibility of Lactobacillus spp. to antimicrobial agents. Int. J. Food Microbiol. 82, 1–11. https://doi.org/10.1016/S0168-1605(02)00254-4
87. Manson, J.M., Keis, S., Smith, J.M.B., Cook, G.M., 2004. Acquired bacitracin resistance in Enterococcus faecalis is mediated by an ABC transporter and a novel regulatory protein, BcrR. Antimicrob. Agents Chemother. 48, 3743–3748. https://doi.org/10.1128/AAC.48.10.3743-3748.2004
88. Ma, X., Wang, H., Deng, Y., Liu, Z., Xu, Y., Pan, X., Musser, J.M., Graviss, E.A., 2006. rpoB gene mutations and molecular characterization of rifampin-resistant Mycobacterium tuberculosis isolates from Shandong Province, China. J. Clin. Microbiol. 44, 3409–3412. https://doi.org/10.1128/JCM.00515-06
89. 89 Casarotti, S.N., Carneiro, B.M., Todorov, S.D., Nero, L.A., Rahal, P., Penna, A.L.B., 2017. In vitro assessment of safety and probiotic potential characteristics of Lactobacillus strains isolated from water buffalo mozzarella cheese. Ann. Microbiol. 67, 289–301. https://doi.org/10.1007/s13213-017-1258-2
90. Vélez, M.P., De Keersmaecker, S.C.J., Vanderleyden, J., 2007. Adherence factors of Lactobacillus in the human gastrointestinal tract. FEMS Microbiol. Lett. 276, 140–148. https://doi.org/10.1111/j.1574-6968.2007.00908.x
91. Hymes, J.P., Johnson, B.R., Barrangou, R., Klaenhammer, T.R., 2016. Functional analysis of an S-layer-associated fibronectin-binding protein in Lactobacillus acidophilus NCFM. Appl. Environ. Microbiol. 82, 2676–2685. https://doi.org/10.1128/AEM.00024-16
92. Muñoz-Provencio, D., Pérez-Martínez, G., Monedero, V., 2010. Characterization of a fibronectin-binding protein from Lactobacillus casei BL23. J. Appl. Microbiol. 108, 1050–1059. https://doi.org/10.1111/j.1365-2672.2009.04508.x

93. Salzillo, M., Vastano, V., Capri, U., Muscariello, L., Sacco, M., Marasco, R., 2015. Identification and characterization of enolase as a collagen-binding protein in Lactobacillus plantarum. J. Basic Microbiol. 55, 890–897. https://doi.org/10.1002/jobm.201400942
94. Díaz-Ramos, À., Roig-Borrellas, A., García-Melero, A., López-Alemany, R., 2012. α-enolase, a multifunctional protein: Its role on pathophysiological situations. J. Biomed. Biotechnol. 2012. https://doi.org/10.1155/2012/156795
95. Fisher, K., Phillips, C., 2009. The ecology, epidemiology and virulence of Enterococcus. Microbiology 155, 1749–1757. https://doi.org/10.1099/mic.0.026385-0
96. Madsen, K.T., Skov, M.N., Gill, S., Kemp, M., 2017. Virulence Factors Associated with Enterococcus Faecalis Infective Endocarditis: A Mini Review. Open Microbiol. J. 11, 1–11. https://doi.org/10.2174/1874285801711010001
97. Chajęcka-Wierzchowska, W., Zadernowska, A., Łaniewska-Trokenheim, Ł., 2017. Virulence factors of Enterococcus spp. presented in food. LWT - Food Sci. Technol. 75, 670–676. https://doi.org/10.1016/j.lwt.2016.10.026
98. Nallapareddy, S.R., Qin, X., Weinstock, G.M., Hook, M., Murray, B.E., 2000. Enterococcus faecalis adhesin, Ace, mediates attachment to extracellular matrix proteins collagen type IV and laminin as well as collagen type I. Infect. Immun. 68, 5218–5224. https://doi.org/10.1128/IAI.68.9.5218-5224.2000
99. Bittencourt De Marques, E.B., Suzart, S., 2004. Occurrence of virulence-associated genes in clinical Enterococcus faecalis strains isolated in Londrina, Brazil. J. Med. Microbiol. 53, 1069–1073. https://doi.org/10.1099/jmm.0.45654-0
100. Hashem, Y.A., Amin, H.M., Essam, T.M., Yassin, A.S., Aziz, R.K., 2017. Biofilm formation in enterococci: Genotype-phenotype correlations and inhibition by vancomycin. Sci. Rep. 7, 1–12. https://doi.org/10.1038/s41598-017-05901-0
101. de Moraes, G.M.D., de Abreu, L.R., do Egito, A.S., Salles, H.O., da Silva, L.M.F., Nero, L.A., Todorov, S.D., dos Santos, K.M.O., 2017. Functional Properties of Lactobacillus mucosae Strains Isolated from Brazilian Goat Milk. Probiotics Antimicrob. Proteins 9, 235–245. https://doi.org/10.1007/s12602-016-9244-8
102. Kafil, H.S., Mobarez, A.M., 2015. Assessment of biofilm formation by enterococci isolates from urinary tract infections with different virulence profiles. J. King Saud Univ. - Sci. 27, 312–317. https://doi.org/10.1016/j.jksus.2014.12.007
103. Ali, L., Goraya, M.U., Arafat, Y., Ajmal, M., Chen, J.L., Yu, D., 2017. Molecular mechanism of quorum-sensing in Enterococcus faecalis: Its role in virulence and therapeutic approaches. Int. J. Mol. Sci. 18. https://doi.org/10.3390/ijms18050960

104. Eaton, T.J., Gasson, M.J., 2001. Molecular Screening of Enterococcus Virulence Determinants and Potential for Genetic Exchange between Food and Medical Isolates. Appl. Environ. Microbiol. 67, 1628–1635. https://doi.org/10.1128/AEM.67.4.1628-1635.2001

105. Cotter, P.D., Hill, C., Ross, R.P., 2005. Food microbiology: Bacteriocins: Developing innate immunity for food. Nat. Rev. Microbiol. 3, 777–788. https://doi.org/10.1038/nrmicro1273

106. Shankar, N., Coburn, P., Pilla, C., Haas, W., Gilmore, M., 2004. Enterococcal cytolysin: Activities and association with other virulence traits in a pathogenicity island. Int. J. Med. Microbiol. 293, 609–618. https://doi.org/10.1078/1438-4221-00301

107. Soleymanzadeh, N., Mirdamadi, S., Kianirad, M., 2017. Incidence of Virulence Determinants and Antibiotic Resistance in Lactic Acid Bacteria Isolated From Iranian Traditional Fermented Camel Milk (Chal). J. Food Biosci. Technol. 7, 1–8.

108. Lopes, M. de F.S., Simões, A.P., Tenreiro, R., Marques, J.J.F., Crespo, M.T.B., 2006. Activity and expression of a virulence factor, gelatinase, in dairy enterococci. Int. J. Food Microbiol. 112, 208–214. https://doi.org/10.1016/j.ijfoodmicro.2006.09.004

109. Todorov, S., Wachsman, M., Ignatova-Ivanova, T., Ivanova, I., 2014. Lactobacillus plantarum st16pa–are we ready to use it as bio-protective culture? Bulg. J. Agric. Sci. 20, 55–58.

110. Holland, T.L., Baddour, L.M., Bayer, A.S., Hoen, B., Miro, J.M., Fowler, V.G., 2017. Infective endocarditis HHS Public Access. Nat. Rev. Dis. Prim. 2, 1–49. https://doi.org/10.1038/nrdp.2016.59.Infective

111. Kirjavainen, P. V., Tuomola, E.M., Crittenden, R.G., Ouwehand, A.C., Harty, D.W.S., Morris, L.F., Rautelin, H., Playne, M.J., Donohue, D.C., Salminen, S.J., 1999. In vitro adhesion and platelet aggregation properties of bacteremia- associated lactobacilli. Infect. Immun. 67, 2653–2655. https://doi.org/10.1128/iai.67.5.2653-2655.1999

112. Esmon, C.T., 1983. Protein-C: biochemistry, physiology, and clinical implications. Blood 62, 1155–1158.

113. Tans, G., Rosing, J., Griffin, J.H., 1983. Sulfatide-dependent autoactivation of human blood coagulation factor XII (Hageman factor). J. Biol. Chem. 258, 8215–8222.

114. Mertens, K., Bertina, R.M., 1980. Pathways in the activation of human coagulation factor X. Biochem. J. 185, 647–658. https://doi.org/10.1042/bj1850647

115. Połka, J., Morelli, L., Patrone, V., 2016. Microbiological cutoff values: A critical issue in phenotypic antibiotic resistance assessment of lactobacilli and bifidobacteria. Microb. Drug Resist. 22, 696–699. https://doi.org/10.1089/mdr.2015.0328

116. Sanders, M.E., Akkermans, L.M.A., Haller, D., Hammerman, C., Heimbach, J., Hörmannsperger, G., Huys, G., Levy, D.D., Lutgendorff, F., Mack, D., Phothirath, P., Solano-Aguilar, G., Vaughan, E., 2010. Safety assessment of probiotics for human use. Gut Microbes 1, 1–22. https://doi.org/10.4161/gmic.1.3.12127

117. Huys, G., D'Haene, K., Swings, J., 2002. Influence of the culture medium on antibiotic susceptibility testing of food-associated lactic acid bacteria with the agar overlay disc diffusion method. Lett. Appl. Microbiol. 34, 402–406. https://doi.org/10.1046/j.1472-765X.2002.01109.x

118. Klare, I., Konstabel, C., Müller-Bertling, S., Reissbrodt, R., Huys, G., Vancanneyt, M., Swings, J., Goossens, H., Witte, W., 2005. Evaluation of new broth media for microdilution antibiotic susceptibility testing of lactobacilli, pediococci, lactococci, and bifidobacteria. Appl. Environ. Microbiol. 71, 8982–8986. https://doi.org/10.1128/AEM.71.12.8982-8986.2005

119. Klare, I., Konstabel, C., Werner, G., Huys, G., Vankerckhoven, V., Kahlmeter, G., Hildebrandt, B., Müller-Bertling, S., Witte, W., Goossens, H., 2007. Antimicrobial susceptibilities of Lactobacillus, Pediococcus and Lactococcus human isolates and cultures intended for probiotic or nutritional use. J. Antimicrob. Chemother. 59, 900–912. https://doi.org/10.1093/jac/dkm035

120. FEEDAP, 2005. Opinion of the Scientific Panel on additives and products or substances used in animal feed (FEEDAP) on the updating of the criteria used in the assessment of bacteria for resistance to antibiotics of human or veterinary importance. EFSA J. 3, 1–12. https://doi.org/10.2903/j.efsa.2005.223

121. European Commission, 2002. Opinion of the Scientific Committee on Ani- mal Nutrition on the criteria for assessing the safety of micro-organisms resistant to antibiotics of human clinical and veterinary importance. Euro- pean Commission, Health and Consumer Protection Directorate G.

122. Liu, B., Zheng, D., Jin, Q., Chen, L., Yang, J., 2019. VFDB 2019: A comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res. 47, D687–D692. https://doi.org/10.1093/nar/gky1080

123. Zhang, Z.Y., Liu, C., Zhu, Y.Z., Wei, Y.X., Tian, F., Zhao, G.P., Guo, X.K., 2012. Safety assessment of Lactobacillus plantarum JDM1 based on the complete genome. Int. J. Food Microbiol. 153, 166–170. https://doi.org/10.1016/j.ijfoodmicro.2011.11.003

124. 124 Li, Shaoying, Li, Z., Wei, W., Ma, C., Song, X., Li, Shufen, He, W., Tian, J., Huo, X., 2015. Association of mutation patterns in GyrA and ParC genes with quinolone resistance

levels in lactic acid bacteria. J. Antibiot. (Tokyo). 68, 81–87. https://doi.org/10.1038/ja.2014.113

125. Rao Thumu, S.C., Halami, P.M., 2012. Presence of erythromycin and tetracycline resistance genes in lactic acid bacteria from fermented foods of Indian origin. Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol. 102, 541–551. https://doi.org/10.1007/s10482-012-9749-4

Chapter 3: WGA-LP: a pipeline for Whole Genome Assembly of contaminated reads

Rossi N. [1], Colautti A. [2], Iacumin L. [2] and Piazza C. [1],

[1]Department of Mathematics, Computer Science, and Physics, University of Udine, Udine, 33100, Italy.
[2]Department of Agricultural, Food, Environmental and Animal Science, University of Udine, via Sondrio 2/A, 33100 Udine, Italy

*Corresponding author:
Nicolò Rossi,
E-mail: olocin.issor@gmail.com

Bioinformatics, 2021, 1–3
https://doi.org/10.1093/bioinformatics/btab719
Advance Access Publication Date: 20 October 2021
Applications Note

Genome Analysis

WGA-LP: a pipeline for whole genome assembly of contaminated reads

N. Rossi [1,*], A. Colautti[2], L. Iacumin[2] and C. Piazza[1]

[1]Department of Mathematics, Computer Science, and Physics, University of Udine, 33100 Udine, Italy and [2]Dipartimento di Scienze Agroalimentari, Ambientali e Animali, University of Udine, 33100 Udine, Italy

*To whom correspondence should be addressed.

Associate Editor: Can Alkan

Received on July 27, 2021; revised on September 22, 2021; editorial decision on October 11, 2021; accepted on October 15, 2021

Abstract

Summary: Whole genome assembly (WGA) of bacterial genomes with short reads is a quite common task as DNA sequencing has become cheaper with the advances of its technology. The process of assembling a genome has no absolute golden standard and it requires to perform a sequence of steps each of which can involve combinations of many different tools. However, the quality of the final assembly is always strongly related to the quality of the input data. With this in mind we built WGA-LP, a package that connects state-of-the-art programs for microbial analysis and novel scripts to check and improve the quality of both samples and resulting assemblies. WGA-LP, with its conservative decontamination approach, has shown to be capable of creating high quality assemblies even in the case of contaminated reads.

Availability and implementation: WGA-LP is available on GitHub (https://github.com/redsnic/WGA-LP) and Docker Hub (https://hub.docker.com/r/redsnic/wgalp). The web app for node visualization is hosted by shinyapps.io (https://redsnic.shinyapps.io/ContigCoverageVisualizer/).

Contact: olocin.issor@gmail.com

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

A currently active challenge in the context of whole genome assembly (WGA) for bacterial genomes is to produce reliable WGAs that are contaminant free (Chun *et al.*, 2018; Steinegger and Salzberg, 2020; Del Angel *et al.*, 2018). In this context, we built WGA-LP, a pipeline that includes different strategies to guide the users in producing higher quality WGAs of prokaryotic genomes, by also including specific features to control possible contamination. Moreover, its workflow is structured to assist in the quality evaluation of the results of each step of the pipeline by providing useful plots and summaries. The current state-of-the-art for decontamination consists in the use of Kraken2 (Wood *et al.*, 2019), a software for read origin imputation, and of pipelines like ProDeGe (Tennessen *et al.*, 2016) and SIDR (Fierst and Murdock, 2017). This last is, however, meant for eukaryotic genomes.

2 Software description

WGA-LP software is built to be used from the command line. The procedures of the pipeline are organized by functionality and have a consistent syntax for argument passing. More details are available in the Supplementary Material, on the GitHub and Docker Hub web pages of the tool.

WGA-LP performs many steps that can be run independently. In order to execute the whole workflow, the user is required to provide the raw reads (.fastq) and, optionally, the references that should be used for decontamination (.fasta). All the other input files can be produced using WGA-LP commands. Check the Supplementary Material for a complete explanation of all the input parameters for WGA-LP.

The first step of WGA-LP has the role of assessing the quality of the input reads and detecting possible contamination sources. To this end, WGA-LP relies on Trimmomatic (Bolger *et al.*, 2014), FastQC (Andrews, 2010), Kraken2 and Bracken (Lu *et al.*, 2017). The trimming step is fully configurable so that the user can choose the right approach for his data.

A novel contribution of WGA-LP is its decontamination procedure, that exploits a custom script including calls to three programs: BWA mem (Li, 2013), Samtools, (Li *et al.*, 2009) and Bazam (Sadedin and Oshlack, 2019). The inputs for the decontamination are the raw reads and two sets of references, one for the target organism and one for the contaminants. We first determine all the reads that map to any contaminant reference, then among such reads we filter the ones that map to any reference genome of the target organism. This gives us the set of reads that we consider to from the contaminant and we remove

3.1 Abstract

Summary: Whole genome assembly (WGA) of bacterial genomes with short reads is a quite common task as DNA sequencing has become cheaper with the advances of its technology. The process of assembling a genome has no absolute golden standard and it requires to perform a sequence of steps each of which can involve combinations of many different tools. However, the quality of the final assembly is always strongly related to the quality of the input data. With this in mind we built WGA-LP, a package that connects state-of-the-art programs for microbial analysis and novel scripts to check and improve the quality of both samples and resulting assemblies. WGA-LP, with its con- servative decontamination approach, has shown to be capable of creating high quality assemblies even in the case of contaminated reads.

Availability and implementation: WGA-LP is available on GitHub (https://github.com/redsnic/WGA-LP) and Docker Hub (https://hub.docker.com/r/redsnic/wgalp). The web app for node visualization is hosted by shinyapps.io (https://redsnic.shinyapps.io/ContigCoverageVisualizer/).

Supplementary information: Supplementary data are available at Bioinformatics online (https://academic.oup.com/bioinformatics/article-abstract/38/3/846/6404579?redirectedFrom=fulltext#supplementary-data).

3.2 Introduction

A currently active challenge in the context of whole genome assembly (WGA) for bacterial genomes is to produce reliable WGAs that are contaminant free (1, 2,3). In this context, we built WGA-LP, a pipeline that includes different strategies to guide the users in producing higher quality WGAs of prokaryotic genomes, by also including specific features to control possible contamination. Moreover, its workflow is structured to assist in the quality evaluation of the results of each step of the pipeline by providing useful plots and summaries. The current state-of-the-art for decontamination consists in the use of Kraken2 (4), a software for read origin imputation, and of pipelines like ProDeGe (5) and SIDR (6). This last is, however, meant for eukaryotic genomes.

3.3 Software description

WGA-LP software is built to be used from the command line. The procedures of the pipeline are organized by functionality and have a consistent syntax for argument passing. More details are available in the Supplementary Material, on the GitHub and Docker Hub web pages of the tool. WGA-LP performs many steps that can be run independently. In order to execute the whole workflow, the user is required to provide the raw reads (.fastq) and, optionally, the references that should be used for decontamination (.fasta). All the other input files can be produced using WGA-LP commands. Check the Supplementary Material for a complete explanation of all the input parameters for WGA-LP. The first step of WGA-LP has the role of assessing the quality of the input reads and detecting possible contamination sources. To this end, WGA-LP relies on Trimmomatic (7), FastQC (8), Kraken2 and Bracken (9). The trimming step is fully configurable so that the user can choose the right approach for his data. A novel contribution of WGA-LP is its decontamination procedure, that exploits a custom script including calls to three programs: BWA mem (10), Samtools, (11) and Bazam (12). The inputs for the decontamination are the raw reads and two sets of references, one for the target organism and one for the contaminants. We first determine all the reads that map to any contaminant reference, then among such reads we filter the ones that map to any reference genome of the target organism. This gives us the set of reads that we consider to from the contaminant and we remove them from the original set. The combination of BWA mem, Samtools (view) and Bazam allows us to simply perform a loop in which fastq reads are mapped to a reference obtaining a bam file. Such file is then processed with Samtools to extract mapped/nonmapped reads. The mapped reads are finally converted back to fastq format through Bazam. The presented decontamination approach is conservative and reduces the probability of discarding reads of the target organism. More details about this approach are presented in Figure 1 and in the Supplementary Materials. This part of the pipeline can be used as a standalone program and can be combined with any other program for

WGA. WGA-LP natively supports SPAdes (13) and Minia (14) assemblers. SPAdes is currently a common choice for bacterial WGA, while Minia is a very simple and fast assembler. The other steps of WGA-LP can support any assembler that includes in its outputs a fasta formatted assembly and a fastg assembly graph [required only for putative plasmid search with Recycler (15)]. We use the term node to refer to an assembled segment of con- tiguous DNA (either a scaffold or a contig) produced by an assembler. WGA-LP includes custom scripts to help in the visualization of node coverage by postprocessing the output of Samtools depth. This allows to produce coverage plots (computed by remapping the reads to the assembled genome) that can be helpful in finding anomalies, such as prophage insertions in the genome. Moreover, WGA-LP provides a web app and tools for nodes (and reads) selection that can improve the decontamination results. These act by exploiting the assembly process as it tends to assemble nodes with reads of the same organism. Such procedures are well fitted to be combined with Kraken2, since this tool can point out problematic nodes, that can be then further evaluated with BLAST alignment (16) in order to validate user selections. For node reordering, WGA-LP uses the ContigOrderer option from Mauve aligner (17). This step requires to provide a reference for the target organism. WGA-LP offers interfaces to two programs that extract putative plasmids: plasmidSPAdes (18) and Recycler. It is highly recommended to check the results of these tools using BLAST. WGA-LP includes three programs to evaluate the quality of the final result of the pipeline: Quast (19) CheckM (20) and Merqury (21). Especially, CheckM is useful to verify the completeness and contamination of the produced assembly. For the annotation, WGA-LP interfaces with Prokka (22) in order to create NCBI compliant assemblies. This can be considered as the final output of the pipeline and can be used for downstream analysis.

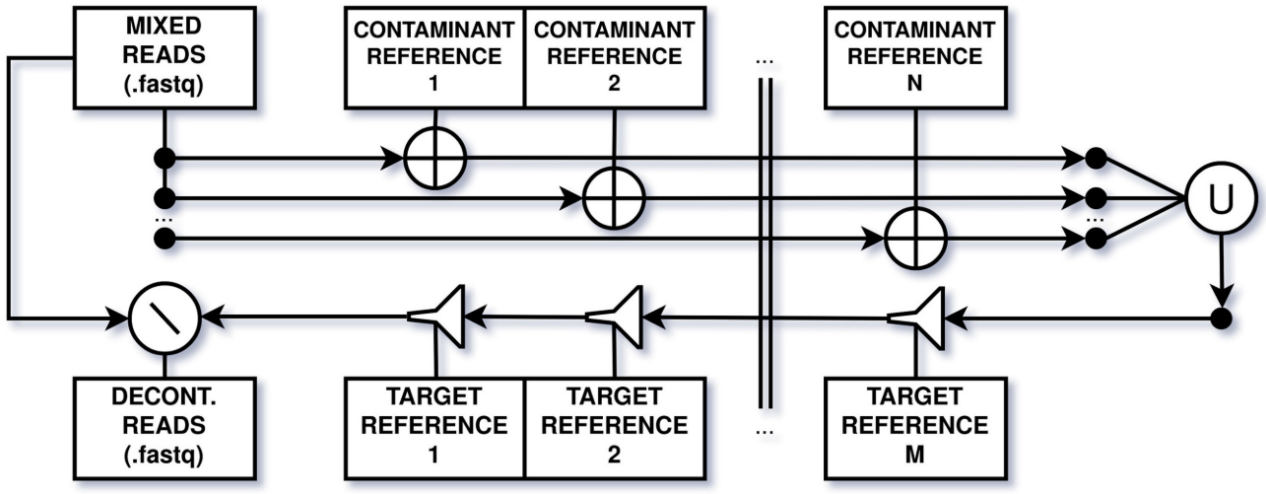


Fig. 1. The decontamination procedure. Input reads are mapped against each reference of the contaminant independently [first three wires from left to are then merged together (Union, [)] and gradually filtered (last wire from right to left), with the effect of removing all the reads that map to any target reference. The final decontaminated reads are extracted by set difference (\) using the original input set

3.4 Results

We tested WGA-LP pipeline on real and simulated data (see Section 4) and we have shown how its workflow was effective in producing a high quality WGA even in the challenging scenario of a contaminated genome, with improvements in comparison with less curated approaches (see Supplementary Material). Finally, we extended the comparison to include ProDeGe, another state-of-the-art decontamination procedure. ProDeGe alone was not able to filter large nodes of the contaminant; however, it was possible to use WGA-LP procedures based on kraken2 classification to refine the resulting assembly, achieving comparable results with our pipeline. However, also in this case, our tool performed better on the elimination of the shorter nodes, keeping those that, in a further check, were classified from the target genome by BLAST alignment. Relying on ART (23), we ran a set of simulations to assess the performance of our decontamination procedure in two different settings. In the first, we investigated the impact of the phylogenetic distance of the contaminant on the effectiveness of our approach, while in the second, we addressed the effect of different contamination levels. In every setting, WGA-LP has proven to be effective in removing the reads of the contaminant while preserving the reads from the target. More details about these simulations can be found in the Supplementary Material. Both the decontamination procedure and the node selection, that are the core of our pipeline, can be integrated in any other pipeline for WGA, in the preprocessing and postprocessing phases.

3.5 Data availability

The testing reads, from the organism Lacticaseibacillus rhamnosus, heavily contaminated with *Pediococcus acidilactici*, are available in the NCBI's Sequence Read Archive at https://www.ncbi.nlm.nih. gov/bioproject/?term=prjna749304 and can be accessed with the accession number SRR15265000, associated to the BioProject PRJNA749304. WGA-LP includes utilities to quickly access all the resources needed to reproduce the tests presented in this paper. All website and links in this paper and in the Supplementary Material were accessed on the July 25, 2021.

Funding: This work was supported by the Ministero dell'Universita` e della Ricerca, Rome, Italy, FIRB n. RBFR107VML. Conflict of Interest: none declared.

3.6 References

1. Chun, J. et al. (2018) Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int. J. Syst. Evol. Microbiol., 68, 461–466.
2. Steinegger, M. and Salzberg,S.L. (2020) Terminating contamination: large-scale search identifies more than 2,000,000 contaminated entries in genbank. Genome Biol., 21, 1–12.
3. Del Angel, V.D. et al. (2018) Ten steps to get started in genome assembly and annotation. F1000Research, 7, 148.
4. Wood, D.E. et al. (2019) Improved metagenomic analysis with kraken 2. Genome Biol., 20, 1–13.
5. Tennessen, K. et al. (2016) ProDeGe: a computational protocol for fully auto- mated decontamination of genomes. ISME J., 10, 269–272.
6. Fierst, J.L. and Murdock, D.A. (2017) Decontaminating eukaryotic genome assemblies with machine learning. BMC Bioinform., 18, 1–16.
7. Bolger, A.M. et al. (2014) Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics, 30, 2114–2120.
8. Andrews, S. (2010) Babraham bioinformatics-FastQC a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham. ac.uk/projects/fastqc (25 July 2021, date last accessed).
9. Lu, J. et al. (2017) Bracken: estimating species abundance in metagenomics data. PeerJ Comput. Sci., 3, e104.
10. Li, H. (2013) Aligning sequence reads, clone sequences and assembly contigs with bwa-mem. arXiv:1303.3997.
11. Li, H. et al. (2009) The sequence alignment/map format and Samtools. Bioinformatics, 25, 2078–2079.
12. Sadedin, S.P. and Oshlack,A. (2019) Bazam: a rapid method for read extrac- tion and realignment of high-throughput sequencing data. Genome Biol., 20, 1–6.
13. Bankevich, A. et al. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol., 19, 455–477.
14. Chikhi, R. and Rizk, G. (2012) Space-efficient and exact de Bruijn graph repre- sentation based on a bloom filter. In: WABI, Lecture Notes in Computer Science, Vol. 7534. Springer, pp. 236–248.

15. Rozov, R. et al. (2017) Recycler: an algorithm for detecting plasmids from de novo assembly graphs. Bioinformatics, 33, 475–482.
16. Altschul, S.F. et al. (1990) Basic local alignment search tool. J. Mol. Biol., 215, 403–410.
17. Rissman, A.I. et al. (2009) Reordering contigs of draft genomes using the Mauve aligner. Bioinformatics, 25, 2071–2073.
18. Antipov, D. et al. (2016) plasmidSPAdes: assembling plasmids from whole gen- ome sequencing data. Bioinformatics, 32, 3380–3387.
19. Gurevich, A. et al. (2013) Quast: quality assessment tool for genome assem- blies. Bioinformatics, 29, 1072–1075.
20. Parks, D.H. et al. (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res., 25, 1043–1055.
21. Rhie, A. et al. (2020) Merqury: reference-free quality, completeness, and phas- ing assessment for genome assemblies. Genome Biol., 21, 1–27.
22. Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. Bioinformatics, 30, 2068–2069.
23. Lu, J. et al. (2017) Bracken: estimating species abundance in metagenomics data. PeerJ Comput. Sci., 3, e104.
24. Huang, W. et al. (2012) Art: a next-generation sequencing read simulator. Bioinformatics, 28, 593–594.

Chapter 4: Draft Genome Sequences of 14 *Lacticaseibacillus* spp. strains, representatives of a collection of 200 strains

Andrea Colautti[a], Nicolò Rossi[b], Carla Piazza[b], Giuseppe Comi[a], Lucilla Iacumin[a] #

[a] Department of Agricultural, Food, Environmental and Animal Science (Di4A), University of Udine, 33100 Udine, Italy

[b] Department of Mathematics, Computer Science, and Physics, University of Udine, 33100 Udine, Italy

4.1 Abstract

Lactobacilli have a fundamental role in the food industry as starters and probiotics, therefore requiring special attention concerning food safety. In this work, 14 strains selected accordingly to their genetic fingerprint and physiologic characteristics are presented as representatives of a collection of 200 strains

4.2 Text

Lactobacilli are ubiquitous Lactic Acid Bacteria and are of particular interest given their significant presence in fermented and non-fermented foods and in the human commensal microbiota. Thanks to their long history of human use and consumption and their use as probiotics, they have been generally recognized as safe (GRAS) (1). However, despite the numerous possible beneficial effects reported in the literature (2), several cases of infections caused by these bacteria have been reported over time in immunodeficient subjects, likely due to the presence of virulence genes and antibiotic-resistance genes (3). As recently stated by EFSA, the whole-genome sequencing constitutes an adequate tool to taxonomically characterize and carry out a risk assessment by verifying the presence of health concern factors in microorganisms intentionally used in the food chain (4). Therefore, this work aimed to provide the genome sequence of highly diverse representative strains to have the opportunity to clarify some specific genetic traits depending on the origin of the strains and associated with the virulence factors in the strictly connected species *Lacticaseibacillus rhamnosus*, *Lacticaseibacillus paracasei, Lacticaseibacillus casei* and *Lacticaseibacillus zeae* recently reclassified (5, 6). These strains were selected as representative from a collection of 200 strains that were characterized using RAPD, Rep-PCR, Sau-PCR, and MLST on stress related genes. Taking to account all the techniques employed, the fingerprint analysis allowed to clusterize the genetic profiles and select these 14 representative strains for each cluster (7,8,9). For the sequencing process, each freeze-dried strain was cultured in MRS broth (Oxoid, Italy) at 30 °C for 48 h. After centrifugation for 5 min at 5000 × *g* the DNA was extracted using the phenol-chloroform method (10) and genomic libraries were constructed employing the TruSeq DNA PCR-Free LT Kit (Illumina, USA) using 2.5 µg of genomic DNA, which were fragmented with a Bioruptor NGS ultrasonicator (Diagenode, USA) followed by size evaluation using Tape Station 2200 (Agilent Technologies). Library samples were loaded into a Flow Cell V3 600 cycles (Illumina, USA) according to the technical support guide. Draft genome sequencing was performed through the genomic platform consisting of a MiSeq (Illumina, UK) following the protocol of the supplier (Illumina, UK). Fastq files of the 250 bp paired-end reads obtained from targeted genome sequencing of the isolated strains were used as input for genome assemblies. The reads were analyzed and

assembled with WGA-LP pipeline (11) using the following tools included in the pipeline with default settings. The reads were trimmed and the Illumina adapters removed using Trimmomatic v0.39 (12), using FastQC v0.11.9 (13) quality check. The presence of any contamination was verified by Kraken2 v2.0.8-b (14). The assembly was made using SPAdes v3.15.2 (15), reordering the resulting scaffolds by aligning them with reference sequences (*L. casei* 12A NZ_CP006690.1, *L. paracasei* ATCC334 NC_008526.1, *L. rhamnosus* GG NC_013198.1) using Mauve v2.4.0 (16). The final quality of the assemblies was then evaluated by CheckM v1.1.3 (17), SamTools v1.10 (18) and Quast v5.0.2 (19). Functional annotation was carried out on the genomes using PGAP 2022-04-14.build6021 (20).

Data availability. Sequences were deposited in GenBank with PRJNA786620 BioProject accession number. Table 1 reports for each sample the taxonomical identification, the isolation source, GenBank accession number, sequencing and assembly statistics, genomes features of strains.

Acknowledgements

This work was funded by Ministero dell'Istruzione, dell'Università e della Ricerca, Rome, Italy, FIRB n. RBFR107VML.

Table 1: Statistics of assemblied genomes

GeneBank accession no.	SRA accession no.	Raw Reads[a]	Strain	Organism name	Source	Provenience[d]	Genome size[a]	Contigs[a]	N50[a]	G+C content (%)[b]	CDS[a]	Completeness (%)[c]
GCA_028878355.1	SRR17145328	720624	LMG 25883	*L. paracasei*	Dairy Products	LMG	3017070	54	99247	46.25	2882	99.46
GCA_028878315.1	SRR17145327	755994	DSM 4905	*L. paracasei*	Human	DSM	3097123	58	137688	46.29	2954	99.46
GCA_028878305.1	SRR17145323	740902	NRRL B-456	*L. paracasei*	Unknown	ARS	3118403	117	102794	46.19	2997	99.46
GCA_028878235.1	SRR17145322	748130	M268	*L. paracasei*	Dairy Products	POT	2730606	140	57940	46.28	2640	99.39
GCA_028878245.1	SRR17145321	850782	O14	*L. rhamnosus*	Dairy Products	POT	2910638	39	283390	46.7	2693	99.46
GCA_028878215.1	SRR17145320	737698	UD2202	*L. zeae*	Dairy Products	UDI	3038780	42	179246	47.97	2778	99.46
GCA_028878255.1	SRR17145319	886860	I2	*L. paracasei*	Sourdough	CAM	2992737	157	48532	46.41	2835	99.46
GCA_028878205.1	SRR17145318	1742298	UD1001	*L. casei*	Human	UDI	3147269	41	276690	47.88	2900	99.46
GCA_028878145.1	SRR17145317	2788788	N1110	*L. rhamnosus*	Human	CAM	3068245	84	119405	46.57	2848	99.46
GCA_028878115.1	SRR17145316	1930166	N202	*L. rhamnosus*	Human	CAM	2882421	59	123389	46.57	2699	99.46
GCA_028878345.1	SRR17217968	1639344	UD193	*L. rhamnosus*	Dairy Products	UDI	3114057	46	196800	46.69	2912	99.46
GCA_028878125.1	SRR17145326	788482	Mo2	*L. rhamnosus*	Human	CAM	2943670	63	119233	46.62	2706	99.46
GCA_028878105.1	SRR17145325	782760	TMW 1.300	*L. paracasei*	Beer	LTM	3178055	136	60472	46.13	3108	99.46
GCA_028878155.1	SRR17145324	932552	DIALYac	*L. paracasei*	Dairy Products	UDI	3037719	95	125133	46.24	2950	99.46

[a] Determined using PGAP
[b] Determined using Quast
[c] Determined using CheckM
[d] Provenience:
LMG: BCCM/LMG, Belgian Co-ordinated Collections of Micro-organisms (BCCM™), Belgium.
DSM: DSM, Deutsche Sämmlung von Mikroorganismen und Zellkülturen, Braunschweig, Germany.
ARS: ARS Culture (NRRL) Collection, United States Department of Agriculture, USA.
POT: Scuola di Scienze Agrarie, Alimentari e Ambientali, Università degli Studi della Basilicata, Potenza, Italy.
UDI: Dipartimento di Scienze degli Alimenti, Università degli studi di Udine, Udine, Italy.
CAM: Dipartimento di Agricoltura, Ambiente e Alimenti, Università degli Studi del Molise, Campobasso, Italy.
LTM: Lehrstuhl für Technische Mikrobiologie, Technische Universität München, Freising, Germany.

4.3 References

1. Adams MR. 1999. Safety of industrial lactic acid bacteria. J Biotechnol 68:171–178.

2. Colautti A, Orecchia E, Comi G, Iacumin L. 2022. Lactobacilli, a Weapon to Counteract Pathogens through the Inhibition of Their Virulence Factors. J. Bacteriol 204:11

3. Colautti A, Arnoldi M, Comi G, Iacumin L. 2021. Antibiotic resistance and virulence factors in lactobacilli: Something to carefully consider. Food Microbiol 103:103934.

4. EFSA. 2021. EFSA statement on the requirements for whole genome sequence analysis of microorganisms intentionally used in the food chain. EFSA J 19.

5. Zheng J, Wittouck S, Salvetti E, Franz CMAP Harris HMB, Mattarelli P, et al. 2020. 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. 70, 2782–2858.

6. Liu DD, Gu TG. 2020. Proposal to reclassify *Lactobacillus zhaodongensis*, *Lactobacillus zeae*, *Lactobacillus argentoratensis* and *Lactobacillus buchneri* subsp. *silagei* as *Lacticaseibacillus zhaodongensis* comb. nov., *Lacticaseibacillus zeae* comb. nov., *Lactiplantibacillus argentoratensis* comb. nov. and *Lentilactobacillus buchneri* subsp. *silagei* comb. nov., respectively and *Apilactobacillus kosoi* as a later heterotypic synonym of *Apilactobacillus micheneri.* Int J Syst Evol Microbiol. 70(12):6414-6417.

7. Iacumin L, Ginaldi F, Manzano M, Anastasi V, Reale A, Zotta T, et al. 2015. High resolution melting analysis (HRM) as a new tool for the identification of species belonging to the Lactobacillus casei group andcomparison with species-specific PCRs and multiplex PCR. Food Microbiol. 46, 357–367.

8. Camprini L, Ginaldi F, Colautti A, Pellegrini M, Comi G, Iacumin L. 2023. Safety traits, genetic and technological characterization of L. casei/paracasei strains. Submitted for publication (under review process).

9. Ginaldi F, Camprini L, Colautti A, Pellegrini M, Comi G, Iacumin L. 2023. Safety traits, genetic and technological characterization of L. rhamnosus strains. Submitted for publication (under review process).

10. Sambrook J, Russel DW. 2001. Molecular Cloning, 3rd ed. Cold Spring Harbor Laboratory Press, New York

11. Rossi N, Colautti A, Iacumin L, Piazza C. 2021. WGA-LP: a pipeline for whole genome assembly of contaminated reads. Bioinformatics 1–3.

12. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

13. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

14. Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. bioRxiv 1–13.

15. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477.

16. Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403.

17. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055.

18. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.

19. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.

20. Zhao Y, Wu J, Yang J, Sun S, Xiao J, Yu J. 2012. PGAP: pan-genomes analysis pipeline. Bioinformatics 28:416–418

Chapter 5: An inter-species comparison on the draft genome of 14 L. casei, L. paracasei and L. rhamnosus strains.

Colautti Andrea[1], Rossi Nicolò[2], Piazza Carla[2], Comi Giuseppe[1], Iacumin Lucilla[1]

[1]Department Agricultural Food Environmental and Animal Science, University of Udine, 33100 Udine, Italy

[2] Department of Mathematics, Computer Science, and Physics, University of Udine, 33100 Udine, Italy

5.1 Abstract

Lacticaseibacillus spp. is a highly studied bacterial group given its importance in terms of practical use in the food industry and as probiotics. Despite numerous studies, distinguishing and classifying bacteria belonging to the former "*L. casei* group" remains a difficult task. Among other characterization techniques, the Whole Genome Sequencing provides a comprehensive overview of the genetic characteristics of the bacterium. This comprehensive understanding of the genome allows us to investigate metabolic and physiological characteristics that may not be highlighted by *in vitro* tests. In this regard, 14 strains belonging to the *L. casei* (1), *L. zeae* (1)*, L. paracasei* (7), *L. rhamnosus* (5) species representative of a collection of 200 strains were sequenced. Given the potential applications of these bacteria, particular attention was paid on the possible presence of virulence and antibiotic resistance traits in their genomes. Plasmids, prophages, CRISPR-Cas systems, IS, and bacteriocins were also detected. Finally, the metabolic pathways that distinguish each strain were compared.

5.2 Introduction

The classification of the "*Lactobacillus casei* group", which until 2020 included the genetically correlated species *L. casei, L. paracasei* and *L. rhamnosus* has been a complicated process that resulted in several reclassifications over time. Starting with the first publication of the comb. "*L. casei*" (1), there were several proposed modifications that even questioned the reference strains (2); (3). These discussions and modifications continue to this day, for example with the current reclassification of *L. zeae* from *L. casei* (4), or the proposed genera *Lacticaseibacillus* ((5), which remain relatively heterogeneous with respect to their AAI values, referring to the casei-group lactobacilli, and including *Lcb. casei, baoquingensis, Lcb. brantae*, *Lcb. cammelliae*, *Lcb. chiayiensis*, *Lcb. hulanensis*, *Lcb. jixianensis*, *Lcb. manihotivorans*, *Lcb. nasuensis*, *Lcb. pantheris*, *Lcb. paracasei (*with two subspecies *tolerans* and *paracasei), Lcb. porcinae*, *Lcb. rhamnosus, Lcb. saniviri*, *Lcb. sharpeae*, *Lcb. songhuajiangensis*, and *Lcb. thailandensis*. Several techniques, such as RAPD and PFGE, have been used in the past to attempt to classify these species, but have failed to produce unambiguous results (6). The 16S rRNA gene sequencing, a widely used golden standard for a rapid classification of many microorganisms is, in this case, not entirely effective. Possible causes of this failure are the high sequence similarity among lactobacilli belonging to the *L. casei* group, as well as the presence of polymorphisms, which can lead to misidentification (7). Several strategies are being investigated in order to effectively differentiate and classify this bacterial group (8). In recent years, different laboratory techniques have been used, such as MALDI-TOF MS (9), multiplex PCR (10), HRM and species-specific PCR (11), comparative analysis of alternative sequences to the 16S

rRNA gene sequence (12)(13)(14), and finally, whole-genome analysis (15)(16). The best characterization strategies, however, appear to be combined ones, with the WGS serving as verification in cases of dubious results. This technique, whose cost and difficulty of implementation have decreased over time, makes the entire genome available for analysis, allowing comparison of many target genes and providing a general overview of the bacterium's metabolic functions. This ability comes in handy when looking for potential probiotics. According to FAO and WHO, in order to be considered a probiotic, a bacterium must be carefully classified at the strain level. In fact, incorrect identification on the label of functional products may be considered food fraud, so *in vitro* and *in vivo* tests must be performed to evaluate potential beneficial or negative factors (17). In this regard, EFSA issued a statement requiring a risk assessment analysis via WGS for all microorganisms intentionally used in the food chain, rather than just probiotics, in order to obtain a correct identification of the used bacteria (18). For taxonomic identification, DNA-DNA hybridization is considered the reference method, but it has the disadvantage of requiring many strains and being time and resources consuming. Currently, several easy-to-use programs that can mimic this *in silico* analysis have been developed. These programs calculate the dDDH using the assembled genome as input, making this type of analysis simpler while still effective (19)(20)(21). In terms of risk assessment, the availability of the entire genome allows for the evaluation of the presence of potential probiotic genes or, conversely, the presence of virulence genes within the microorganism. This approach can help researchers by suggesting possible targets for *in vitro* and *in vivo* experiments as it takes into consideration also partial, or unexpressed genes that may not be observed otherwise. However, all these analyses presuppose a correct and careful assembly to mitigate the increase in the amount of errors that are identified in deposited sequences (22)(19). Very often, the genomes are directly assembled without an initial check to diagnose for any contamination in the reads or sequencing errors, and often final quality checks of the assembled genome are missing. For this reason, it was decided to opt for a pipeline that would allow both the identification of contaminations and the detection of assembly errors. In this article, 14 genomes belonging to *L. casei, L. zeae, L. paracasei*, and *L. rhamnosus* were selected, identified and characterized from a previous work (11) that considered a library of about 200 strains. The aspects considered for this selection were the genetic and phenotypic characterization of the bacteria on the basis of different characteristics demonstrated and analysed *in silico*, the verification of their correct identification through ANI and dDDH, the characterization of their metabolisms in comparison with the deposited references, and the presence of any genetic characteristics for possible use in the food and probiotic fields, with a particular focus on the presence of virulence factors.

5.3 Materials and methods

Genomes

The 14 (1 *L. casei*, 1 *L. zeae,* 7 of *L. paracasei*, and 5 *L. rhamnosus*) described in the previous chapter "Draft Genome Sequences of 14 *Lacticaseibacillus* spp., representatives of a collection of 200 strains" were previously identified by species-specific PCRs, multiplex PCR, and High resolution melting analysis (11). These strains were isolated from different matrices such as cheese (strains LMG, 25883, M268, O14, UD2202, DIALYAC) human tissues/secretions (strains DSM4905, UD1001, N1110, N202, Mo2) sourdoughs (strain I2) spoiled wine (strain UD193) beer (strain TMW 1.300) and unknown sources (strain NRRL B-456), and selected for genome sequencing based on RAPD, REP, SAU profiles as representative of specific clusters of the stock library consisting of 200 strains (data not shown). To control the assembly process, genomes were processed with the WGA-LP pipeline described in (23) using SPAdes (24) as the assembler. In addition, part of the following characterization tools were used through this pipeline with default settings.

Genome analysis

The general features and quality assessment of the assembled genomes were computed using Quast (25), and CheckM (26), while Recycler was used to identify plasmids (27). Further verifications on predicted plasmidic sequences were made by aligning the sequences on the nucleotide BLAST suite and by evaluating the correspondence to plasmids or chromosomal portions. Digital-DNA/DNA hybridization (dDDH) was calculated using TYGS (28), and the average nucleotide identity (ANI) was calculated using ANI Matrix (29), setting the demarcation limits to distinguish two different species as 95% for the ANI and 70% for the dDDH (30). IS sequence prediction was performed using the ISfinder tool (31). The presence of resistance factors was verified through the Resistance Gene Identifier tool (RGI) from the Comprehensive Antibiotic Resistance Database (CARD) (32) and ResFinder 4.0 using acquired antimicrobial resistance genes search function (33). In addition, the antibiotic resistance and virulence genes, whose presence has been reported in the literature for Lactobacilli (34) together with genes with probiotic effect reported in the literature for lactobacilli (35)(36)(37), shown in **Table 1**, were searched through BLAST and reported with at least 70 % of query cover and 80 % of percent identity (18). The presence of prophages was investigated by PHASTER (38) and Prophage Hunter (39). CRISPR-Cas were searched with the CRISPRFinder tool (40).

Table 1, Antibiotic resistance, virulence and probiotic genes searched

Antibiotic Resistance									
tet(M)	AAA24784.1	*msrC*	AAK01167.1	*aac(6')-Ie-aph(2")-Ia*	AHY23917.1	*van*X	ADM24921.1	*cat*	WP_110139844.1
tet(K)	AXY65082.1	*lnu*(A)	ACC61208.1	*aph(3")-III*	AQY75633.1	*bla*CTX-M	AEL20750.1	*cat*A	ODO61704.1
tet(L)	AXH80272.1	*vat*(E)	NP_783842.1	*aadA*	QID24729.1	*bla*SHV	ABN49114.1	*cat*-TC	AAB53259.1
erm(A)	QBA99766.1	*aac(6')-Ii*	AAB63533.1	*aadE*	QDD71360.1	*bla*TEM	TXG03870.1	*bcrA*	AJF17087.1
erm(C)	QBC82934.1	*ant(4')-Ia*	QBC83248.1	*ant(6)*	TQA74313.1	*bal*Z	AAA24777.1	*bcrB*	AJF17088.1
mefA	QIS77055.1	*aph(2")-Ib*	AAG13458.1	*vanA*	AAM77885.1	*bla*OXA-1	AVE16060.1	*bcrD*	AJF17089.1
mrsA	AAZ32815.1	*aph(2")-Ic*	QBC83246.1	*vanH*	WP_010815296.1	*bla*OXA-48	AXE72493.1	*bcrR*	AAS78452.1
Potential Virulence									
ace	AAD43342.1	*ccf*	APU94149.1	*cylR2*	AAL60140.1	*cylB*	AAA03343.1	*fsrA*	ACO94083.1
asa1	BAJ34847.1	*cob*	VTS87224.1	*cylLL*	AAA62648.1	*cylA*	AAK67268.1	*fsrB*	ACO94085.1
esp	AAD09858.1	*cpd*	APU51069.1	*cylLs*	EET97079.1	*cylI*	AAM21178.1	*fsrC*	ACO94086.1
efaA	AAO81809.1	*cylR1*	AAL60139.1	*cylM*	AAK67266.1	*gelE*	ACO94084.1	*hyl*	AAN34803.1
Probiotic									
abpT	ABE00714.1	*cps1F*	BAG84623.1	*LBA0867*	AAV42726.1	*LJ1035*	AAS08857.1	*msmE*	AAV42385.1
bfra	AAV42388.1	*cps1G*	BAG84624.1	*LBA0995*	AAV42845.1	*LJ1147*	AAS08969.1	*msrB*	WP_003602400.1
bsh1	CCC80500.1	*cps1H*	BAG84625.1	*LBA0996*	AAV42846.1	*LJ1413*	AAS09179.1	*mub*	AAV43464.1
bshA	AAV42751.1	*cps1I*	BAG84626.1	*lba1272*	AAV43103.1	*LJ1680*	AAS09453.1	*prtP*	WP_003567119.1
bshB	AAV42923.1	*cps1J*	BAG84627.1	*LBA1427*	AAV43252.1	*lp 1403*	CCC78746.1	*prtP1*	CDG41976.1
cdpA	AAV42118.1	*dltA*	WP_047107405.1	*LBA1428*	AAV43253.1	*lp 1403*	CCC78746.1	*prtR*	CAD43138.1
clpC	CAQ67748.1	*dltD*	AAB17660.1	*LBA1429*	AAV43254.1	*lp 2940*	CCC80013.1	*pts14C*	CCC78553.1
clpE	CCK22892.1	*dps*	AEI57033.1	*LBA1430*	AAV43255.1	*lp_2940*	CCC80013.1	*rrp-1*	AAD10258.1
clpL	CAQ67279.1	*fbpA*	CAQ66743.1	*LBA1431*	AAV43256.1	*Lr1265*	ABS84230.1	*rrp-48*	AAD10267.1
copA	CCC80105.1	*fosE*	ABD57319.1	*LBA1432*	AAV43257.1	*lr1516*	ABB02575.1	*slpA*	AAV42070.1
cps1A	BAG84618.1	*gadC*	AAV41961.1	*LBA1524*	AAV43343.1	*Lr1584*	ABS84214.1	*srtA*	CAR88038.1
cps1B	BAG84619.1	*gtfA*	WP_019251826.1	*LBA1663-1664*	AAV43475.1	*lspA*	CCK22598.1	*treC*	AAV42863.1
cps1C	BAG84620.1	*inu*	CAL25302.1	*LJ0056*	AAS08038.1	*luxS*	ABC59818.1	*wzb*	ABV54212.1
cps1D	BAG84621.1	*labT*	AJP47088.1	*LJ1021*	AAS08843.1	*met*	CUU11694.1	*xylA*	WP_194957996.1
cps1E	BAG84622.1	*lamA*	CCC80542.1	*LJ1021*	AAS08843.1	*msa*	CCC78612.1		

Genome annotation and characterization

Genomes were annotated with Prokka (41), and annotations were used to create the gene presence/absence matrix and individuate the core, shell, cloud, and pangenome (42). The obtained pangenome was compared with the pangenome of the references already deposited shown in (**Supplementary Table S1**). Thanks to EggNog Mapper (43), using the amino acid sequences identified by Prokka as input, the COG annotation, and the KEGG annotation were performed. From the obtained KEGG annotation, using the KEGG-Mapper function (44), the functional predictions of the genomes were made, and the metabolic pathways present in the lactobacilli were verified.

5.4 Results

General features

The average total size of the 14 assembled genomes was 3.01 Mbp, varying between 2.73 Mbp of strain M268 and 3.17 Mbp of strain TMW 1.300 with an average %GC of 46.6 %, between 46.13 % of strain TMW 1.300 and 47.97 % of strain UD2202. Genomes were reconstructed from an average of 81 scaffolds, ranging from 40 to 158, of which 3-19 scaffolds made up 50 % of the genome (mean L50 = 9.28, mean N50 = 137854.2). The functional annotation detected an average of 2921 genes present in 2863 coding sequences (CDS), ranging from 2600 of strain M268 to 3061 of strain TMW 1.300, with an average of 56 tRNAs, and 1 tmRNA (**Table 2).**

Investigating the 14 genomes with the help of Roary, a core genome consisting of 353 coding sequences (CDS), a shell genome of 4716 CDS, and a cloud genome of 6160 CDS were obtained from an overall pan-genome of 11229 CDS. From the genes identified, through the COG annotation (**Figure 1**), it can be observed how distributed and conserved are the various proteic functions between the different genomes. The most represented functions are those inherent to the carbohydrate transport and metabolism (13.7%) (G), transcription (11.2%) (K) replication, recombination, and repair (10.5%) (L), and cell wall, membrane, and envelope biogenesis (7.9%) (M), and amino acid transport and metabolism (5.9%) (E). Most of the genes shared in the core genome fall into category J of the COG annotation, which includes translation, ribosomal structure, and biogenesis. As for the shell genome, it is mainly composed of transcription proteins (K), carbohydrate transport and metabolism (G), and amino acid transport and metabolism (E), while the cloud genome was still characterized by a high presence of G and K related functions together with replication, recombination, and repair (L).

Table 2, genomes general features

	LMG 25883	DSM4905	NRRL B-456	M268	O14	UD2202	I2	UD1001	N1110	N202	UD193	Mo2	TMW 1.300	DIALYac
Nodes	59	58	117	140	40	42	158	41	85	59	46	63	137	95
Lenght**	3,02	3,1	3,12	2,73	2,91	3,04	2,92	3,15	3,07	2,88	3,11	2,94	3,18	3,04
%GC *	46,25	46,29	46,19	46,28	46,7	47,97	46,41	47,88	46,57	46,57	46,69	46,62	46,13	46,24
N50 *	99247	137688	102794	57940	283390	179246	48532	276690	119405	123389	196800	119233	60472	125133
N75 *	60717	84986	53038	26289	109666	106862	27464	123180	74376	65295	105793	66192	29788	62858
L50 *	10	7	10	16	3	6	19	6	8	8	5	9	15	8
L75 *	20	14	22	31	7	11	38	9	16	16	10	18	33	16
CDS	2890	2963	2949	2600	2742	2827	2804	2934	2878	2771	2963	2766	3061	2944
Genes	2947	3021	3006	2658	2799	2884	2862	2991	2936	2828	3021	2823	3118	3001
tRNA	56	57	56	57	56	56	57	56	57	56	57	56	56	56
tmRNA	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CRISPR	1	0	2	0	1	3	0	0	1	0	0	1	1	1

*= calculated on Nodes >500bp **=in Mbp

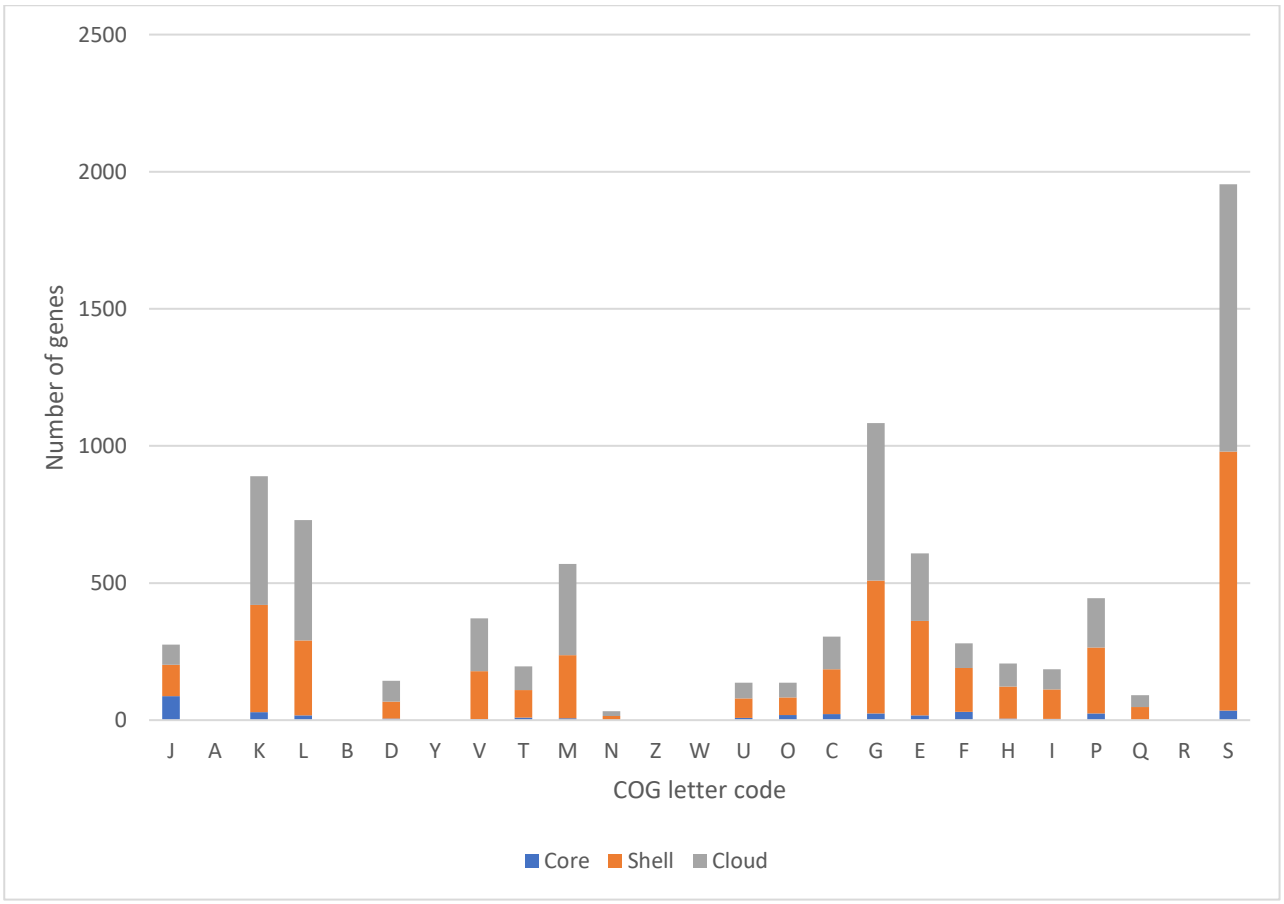


Figure 1, pangenome COG functions annotation

Table 3, any matrix of considered strains against *Lacticaseibacillus* spp. reference strains

L. jixianensis 1594	L. camelliae DSM 22697	L. pantheris DSM 15945	L. songhuajiangensis 7/19	L. baoquingensis 473	L. thailandensis DSM22698	L. nasuensis JCM 19617	L. porcinae JCM19617	L. manihotivorans DSM 13343	L. sanviri DSM 24301	L. sharpeae DSM 20505	L. hulanensis ZW163	M268	L. paracasei subsp tolerans DSM 20258	DIALYac	DSM4905	LMG 25883	I2	L. paracasei DSM 5622	TMW 1.300	NRRL B-456	UD1001	L. casei DSM 20011	UD2202	L. zeae DSM 20178	L. chiayiensis NCYUAS	N202	Mo2	O14	UD193	N1110	L. rhamnosus DSM 20021	L. brantae DSM 23927	
100	82	79	82	80	80	77	79	84	76	77	78	78	78	78	78	78	80	77	78	79	81	80	77	0	81	79	78	81	81	78	79	0	*L. jixianensis 159-4*
82	100	82	82	82	81	77	79	78	79	77	78	78	78	79	79	79	81	78	79	79	82	83	79	0	84	79	78	84	84	77	79	0	*L. camelliae DSM 22697*
79	82	100	78	79	81	77	80	79	77	78	81	78	78	78	78	79	79	79	77	78	78	78	77	0	84	77	78	78	84	78	77	0	*L. pantheris DSM 15945*
82	82	78	100	85	80	75	84	81	78	81	81	78	77	79	77	77	83	78	78	77	84	82	76	0	84	77	77	82	84	78	78	0	*L. songhuajiangensis 7/19*
80	82	79	85	100	84	78	79	80	77	83	81	77	79	78	77	78	81	79	78	80	85	81	78	0	82	78	78	82	82	77	77	0	*L. baoquingensis 47-3*
80	81	81	80	84	100	77	78	82	77	77	78	81	80	78	78	80	84	80	80	79	83	83	78	0	85	78	78	83	84	79	78	0	*L. thailandensis DSM22698*
77	77	77	75	78	77	100	77	78	76	76	76	77	78	77	77	77	77	79	77	77	76	78	77	0	78	77	77	77	76	77	77	0	*L. nasuensis JCM 19617*
79	79	80	84	79	78	77	100	85	78	81	81	77	77	78	78	77	79	85	78	80	76	77	76	0	77	77	78	78	77	77	79	0	*L. porcinae JCM19617*
84	78	79	81	80	82	78	85	100	82	81	78	79	79	78	80	80	83	84	81	82	82	80	78	0	81	80	79	82	81	79	79	0	*L. manihotivorans DSM 13343*
76	79	77	78	77	77	76	78	82	100	77	77	79	78	79	79	78	78	77	79	79	78	78	77	0	77	78	78	78	79	78	78	0	*L. sanviri DSM 24301*
77	77	78	81	83	77	76	81	81	77	100	81	78	78	78	78	77	80	81	79	77	76	76	75	0	78	77	78	77	76	77	82	0	*L. sharpeae DSM 20505*
78	78	81	81	81	78	76	81	78	77	81	100	78	78	78	78	78	78	78	78	78	77	76	77	0	79	77	77	78	77	77	80	0	*L. hulanensis ZW163*
78	78	78	78	77	81	77	77	79	79	78	78	100	99	99	99	99	99	99	98	99	80	80	79	79	79	80	80	80	80	80	80	0	**M268**
78	78	78	77	79	80	78	77	79	78	78	78	99	100	98	98	98	98	98	98	98	81	81	80	79	79	81	80	79	80	80	79	0	*L. paracasei subsp tolerans DSM 20258*
78	79	78	79	78	78	77	78	78	79	78	78	99	98	100	99	99	99	99	99	99	82	81	79	79	79	80	80	80	80	80	80	0	**DIALYac**
78	79	78	77	77	78	77	78	80	79	78	78	99	98	99	100	99	99	99	99	99	81	81	79	79	79	80	80	80	82	80	79	0	**DSM4905**
78	79	79	77	78	80	77	77	80	78	77	78	99	98	99	99	100	99	99	99	99	82	81	80	79	79	81	80	80	81	80	80	0	**LMG 25883**
80	81	79	83	81	84	77	79	83	78	80	78	99	98	99	99	99	100	100	99	99	81	81	80	79	80	81	80	80	81	80	79	0	**I2**
77	78	79	78	79	80	79	85	84	77	81	78	99	98	99	99	99	100	100	99	99	80	80	80	79	80	81	80	79	81	80	80	0	*L. paracasei DSM 5622*
78	79	77	78	78	80	77	78	81	79	79	78	98	98	99	99	99	99	99	100	99	81	81	80	79	80	80	80	80	82	80	80	0	**TMW 1.300**
79	79	78	77	80	79	77	80	82	79	77	78	99	98	99	99	99	99	99	99	100	81	81	79	79	79	80	80	80	81	80	80	0	**NRRL B-456**
81	82	78	84	85	83	76	76	82	78	76	77	80	81	82	81	82	81	80	81	81	100	95	96	94	88	81	81	81	82	81	80	0	**UD1001**
80	83	78	82	81	83	78	77	80	78	76	76	80	81	81	81	81	81	80	81	81	95	100	94	94	88	81	81	81	82	81	81	0	*L. casei DSM 20011*
77	79	77	76	78	78	77	76	78	77	75	77	79	80	79	79	80	80	80	80	79	96	94	100	96	88	81	81	81	81	81	81	0	**UD2202**
0	0	0	0	0	0	0	0	0	0	0	0	79	79	79	79	79	79	79	79	79	94	94	96	100	87	81	81	81	81	81	81	0	*L. zeae DSM 20178*
81	84	84	82	85	78	77	81	77	78	79	79	79	79	79	79	80	80	80	80	79	88	88	88	87	100	81	80	80	81	81	81	0	*L. chiayiensis NCYUAS*
79	79	77	77	78	78	77	77	80	78	77	77	80	81	80	80	81	81	81	80	80	81	81	81	81	81	100	98	98	97	97	98	0	**N202**
78	78	78	77	78	78	77	78	79	78	78	77	80	80	80	80	80	80	80	80	80	81	81	81	81	80	98	100	98	97	97	97	0	**Mo2**
81	84	78	82	82	83	77	78	82	78	77	78	80	79	80	80	80	80	79	80	80	81	81	81	81	80	98	98	100	98	97	98	0	**O14**
81	84	84	82	84	76	77	81	79	76	77	80	80	80	80	82	81	81	81	82	81	82	82	81	81	81	97	97	98	100	100	100	0	**UD193**
78	77	78	78	77	79	77	77	79	78	77	77	80	80	80	80	80	80	80	80	80	81	81	81	81	81	97	97	97	100	100	99	0	**N1110**
79	79	77	78	77	78	77	79	79	78	82	80	80	79	80	79	80	79	80	80	80	80	81	81	81	81	98	97	98	100	99	100	0	*L. rhamnosus DSM 20021*
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	*L. brantae DSM 23927*

Identification

From the analysis of ANI within the Lacticaseibacillus group described by (Zheng et al., 2020), it can be observed how the 14 strains clustered with the Lacticaseibacillus reference strains (Table 3). Within the clusters of Lcb. rhamnosus and Lcb. paracasei, all the analysed strains matched with the reference strains DSM 20021 and DSM5622 with ANI values above 97 %. As for strains UD1001 and UD2202, they clustered with borderline values with Lcb. casei DSM 20011 and Lcb. zeae DSM 20178, however, above demarcation threshold.

Further information on phylogenetic proximity can be obtained from dDDH analysis (Figure 2). In fact, also in this case *Lcb. rhamnosus* and *Lcb. paracasei* strains clustered with the references, while UD1001 clustered between *Lcb. casei* DSM 20011 and strain UD2202, that resulted phylogenetically closer to *Lcb. zeae* KCTC3804, whose distinction as a separate species continues to be debated (5)(4). The search for glycosyltransferase family 8 (KRK10099.1) was carried out as a further discriminating element for the identification of the species *Lcb. zeae* as suggested by Kim et al., 2021. The gene was

present with a similarity of 294/299 identities in strain UD2202 and 292/299 in strain UD1001. However, this gene was also present with high similarity in *Lcb. casei* strains (FBL6 298/299 and N 294/299 identities) and absent in some strains classified as *Lcb. zeae* (MGYG-HGUT-02383, CRBIP24.44, DSM 20178) thus making this method in our opinion ineffective in discriminating between the two species (Supplementary table S2).

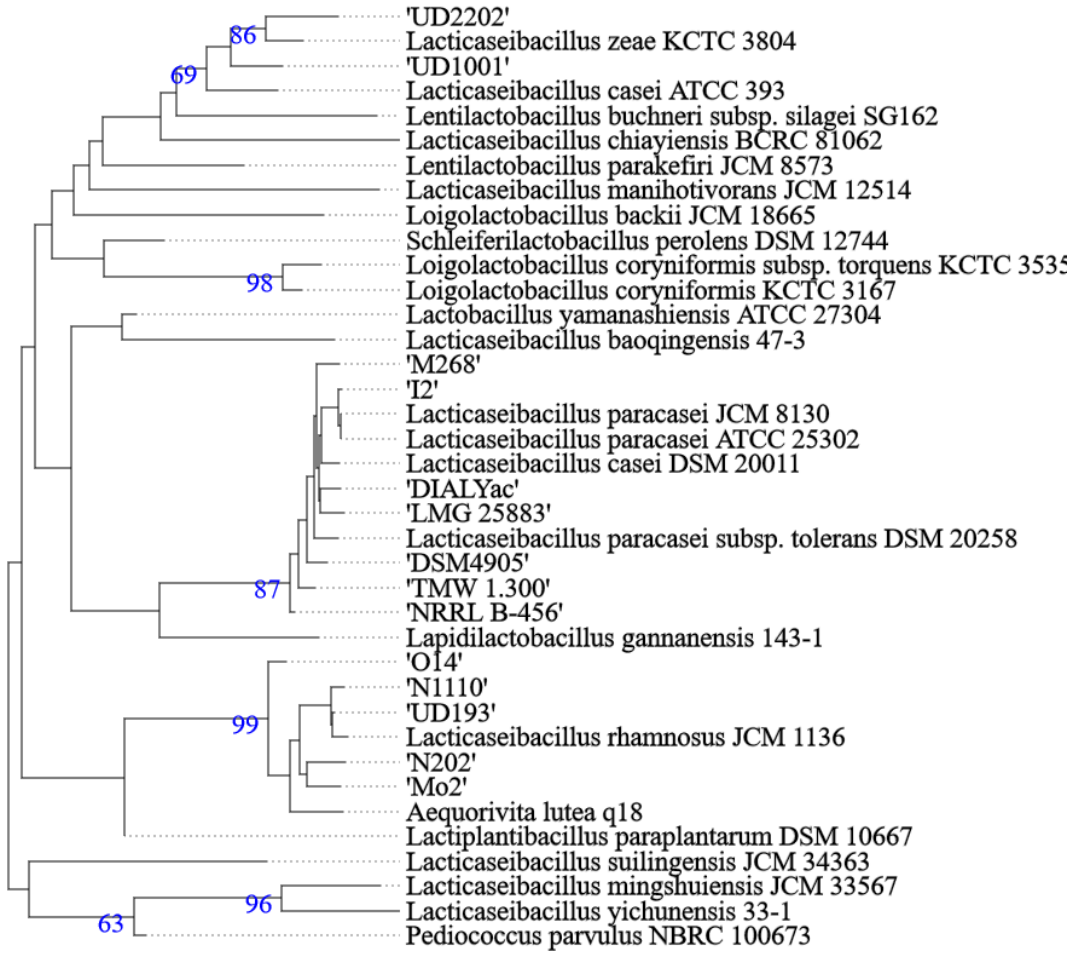


Figure 2, TYGS clusterization of the strains

To get a more complete overview of the genetic characteristics, tabulating the presence/absence of the genes of the *Lcb. casei* and *Lcb. zeae* strains currently deposited on NCBI using Roary, it was possible to observe divisions into specific clusters (Figure 3). A cluster composed by strains 12A, UW4, Z11, A2_362, NBRC101979, UW1, GCRL163, and MJA12 clearly separated, probably being misidentified *Lcb. paracasei* strains deposited as *Lcb. casei*. On the other hand, it can be observed that from the cluster containing the type strain *Lcb. casei* DSM 20011, two groups separate, one containing the strains currently classified as *Lcb. zeae* (CECT9104 MGYG-HGUT-02383

CRBIP2444 DSM20178 CRBIP2458 and FBL8), which also contain strains UD2202 and UD1001 together with some *Lcb. casei* (N87 and B900021), and another cluster consisting of 5 strains currently classified as *Lcb. casei* (BIO5773, LC5, N, 867_LCAS and FBL6). The differences between these clusters, however, were minor, especially when compared to the first cluster, or the closely related species *Lcb. paracasei* and *Lcb. rhamnosus.*

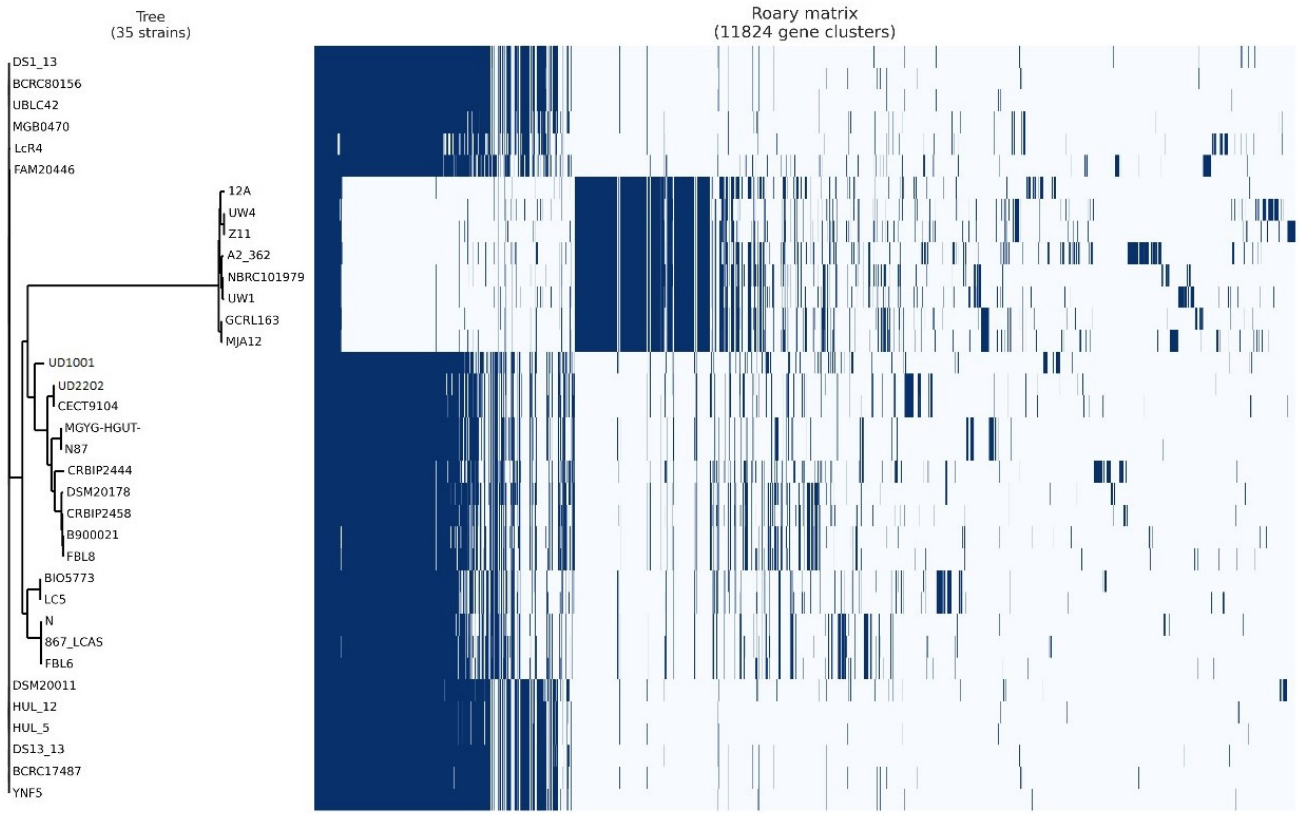


Figure 3, genome clusterization of *Lcb. zeae* and *Lcb. casei* strains made with Roary

Pan-genomes metabolic pathways prediction

Through the KEGG mapper tool, from the annotated genomes (**Figure 4**) the metabolic pathways present in the pangenomes of the considered *Lcb. paracasei,* and *Lcb. rhamnosus* strains obtained through Roary, and from the genomes of *Lcb. casei, Lcb. zeae* strains, have been reconstructed (Supplementary material Table S3). The complete metabolic pathways shared between all species were mainly those related to carbohydrate metabolism (glycolysis, pyruvate oxidation, gluconeogenesis, pentose phosphate cycle, degradation of galactose and ascorbate, biosynthesis and degradation of glycogen, biosynthesis of glucose and galactose, UDP-acetyl-D-glucosamine). Furthermore, *Lcb. rhamnosus* together with *Lcb. zeae* presented the complete pentose phosphate cycle and the transformation of fructose 6P into ribose 5P, which is incomplete by one step in *Lcb. casei* and *Lcb. paracasei*. *Lcb. casei, Lcb. zeae,* and *Lcb. rhamnosus* showed the degradative pathway from

D-glucuronate to pyruvate + D-glyceraldehyde 3P, that is incomplete in *Lcb. paracasei*. A metabolism difference in *Lcb. zeae* was identified in the incomplete degradation of ascorbate pathway, which was complete for the other species.

In energy metabolism, the pathways of transformation of acetyl-CoA into acetate, the assimilation of formaldehyde through ribulose monophosphate, and the synthesis of ATP through F-type ATPase were shared between all species, while *Lcb. paracasei* pan-genome showed the pathway of formaldehyde assimilation through xylulose incomplete only of one part, as for the pathway of transformation of ribulose-5P into glyceraldehyde-3P in *Lcb. rhamnosus*.

In lipid metabolism, the four species shared the biosynthetic pathway of initiation and elongation of fatty acids, while only *Lcb. casei*, *Lcb. Zeae,* and *Lcb. paracasei* demonstrated the capability to synthesize acyl-CoA.

In the metabolism of nucleotides, the synthesis of inosine monophosphate, adenine and guanine, pyrimidine ribonucleotide, and deoxyribonucleotide were shared between all four, while the biosynthesis of uridine monophosphate was incomplete in *Lcb. paracasei* and *Lcb. rhamnosus*.

Considering the amino acid metabolism, the four species shared the threonine, lysine, proline, and histidine biosynthesis. The strains belonging to *Lcb. rhamnosus* showed a complete cysteine pathway both starting from methionine and together with *Lcb. paracasei* starting from serine, also sharing a complete tryptophan synthesis pathway, that was completely absent in *Lcb. casei* and *Lcb. zeae*.

Moreover, all the strains, independently from the species, showed a complete thiamine salvage, CoA, C1-unit interconversion pathway, while *Lcb. casei* strains showed an almost complete tetrahydrofolate synthesis pathway.

As for the biosynthesis capacity of terpenoids and polyketides, all the pangenomes shared the biosynthesis capacity of the C10-C20 isoprenoids, while only *Lcb. rhamnosus* and *Lcb. paracasei* possess the dTDP-L-rhamnose biosynthesis pathway.

As regards the possible resistance to drugs, the presence of a complete NorB in *Lcb. rhamnosus* and *Lcb. casei* strains, and AbcA in *Lcb. paracasei* efflux pump systems are suggested.

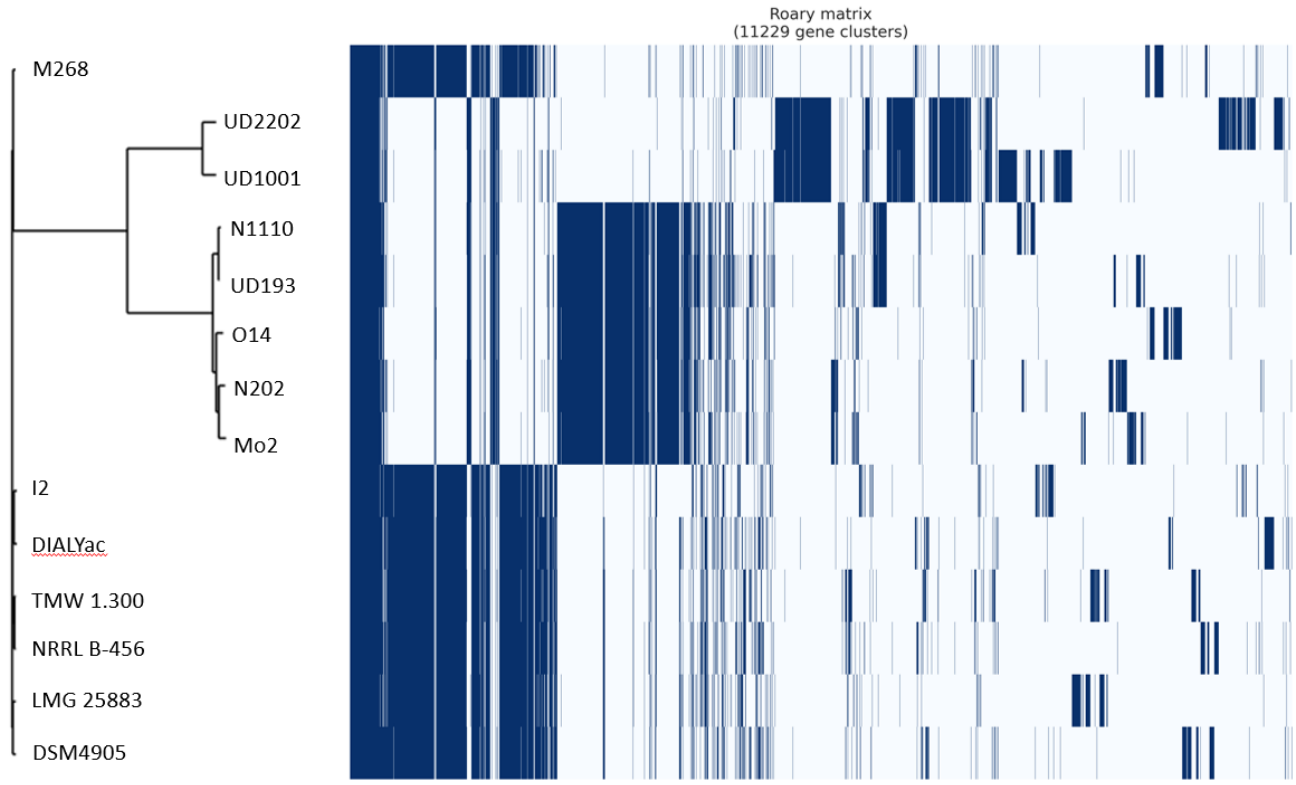


Figure 4, Roary clusterization of the annotated genomes

Mobilome

The presence of IS was investigated in all genomes. Considering the matches with e-value = 0 and bitscore >100, IS belonging to the IS3, ISL3, and IS5 families were found in *L. casei*, IS6 in *L. zeae*, IS3, ISL3, IS30, IS5 in *L. paracasei*, ISLre2, IS5, IS30, IS256, ISL3 in *L. rhamnosus*.

Table 4, IS families predicted with IS finder

	IS Family						
Strain	IS3	IS5	IS6	IS256	IS30	ISL3	ISLre2
LMG 25883	2	-	-	-	2	1	-
DSM4905	2	1	-	-	3	-	-
NRRL B-456	2	2	-	-	2	1	-
M268	2	3	-	-	2	-	-
O14	1	2	-	-	2	-	1
UD2202	-	-	2	-	-	-	-
I2	1	3	-	1	3	-	-
UD1001	1	2	-	-	-	1	-
N1110	1	2	-	1	-	-	-
N202	2	3	-	-	2	1	-
UD193	2	2	-	1	3	1	-
Mo2	1	2	-	-	-	-	1
TMW 1.300	2	2	-	-	3	-	-
DIALYac	2	-	-	-	4	1	-

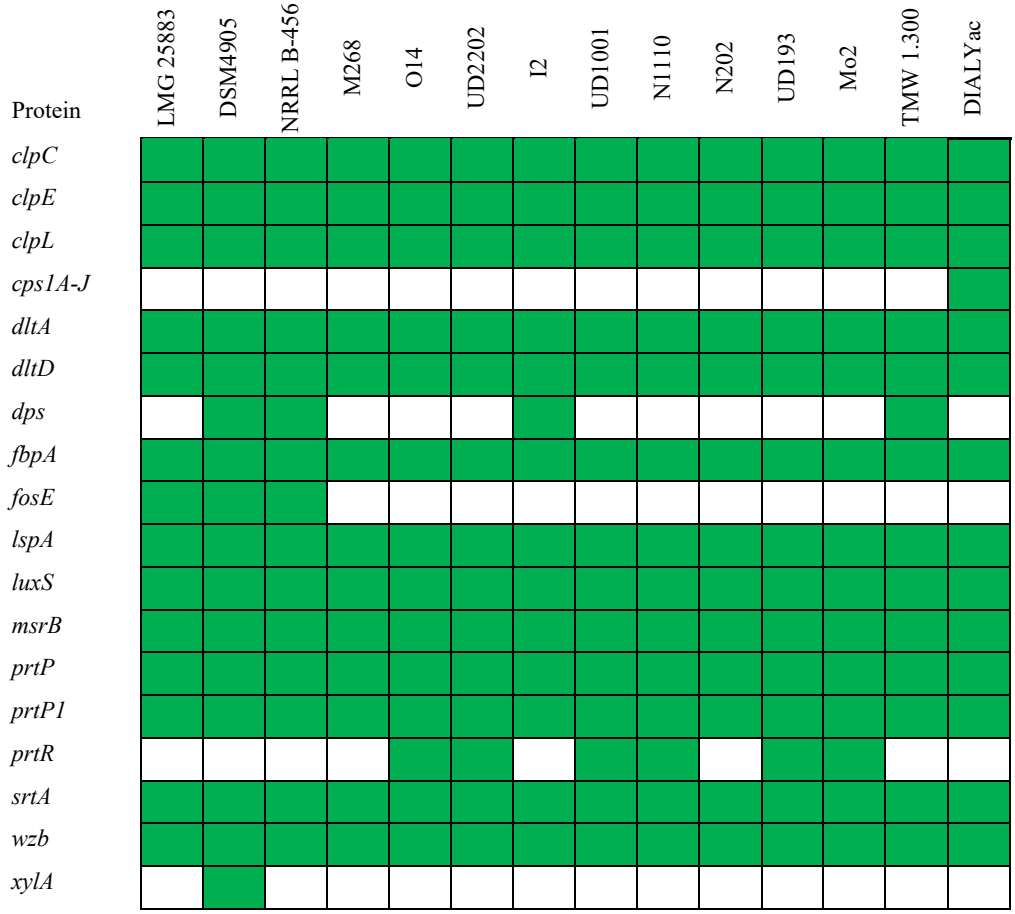
Antibiotic Resistance

Neither RGI nor ResFinder databases detected any Perfect or Strict Hits with antibiotic-resistance genes in the database. Also, from the careful alignment on blast of the virulence and antibiotic resistance genes reported in **Table 1** no match emerged.

Probiotic genes

The probiotic genes identified were resumed in **Table 5**. All strains presented the *clpC, clpE, clpL* ATPase chaperones, *dltA, dltD* genes responsible for the d-Alanylation of LTA, *fbpA* genes that increases adherence, lipoprotein signal peptidase *lspA* gene*, luxS* gene involved in AI-2 and AI -3 metabolism, methionine sulfoxide reductase *msrB*, *prtP* and *prtP1* proteinases, *srtA* sortase*, wzb* involved in the exopolysaccharides biosynthesis, while the *dps* gene involved in stress protection was present in strains DSM4905, NRRL B-456, I2, TMW 1.300, *fosE* gene linked to fructo-oligosaccharides metabolism in strains LMG 25883, DSM4905, and NRRL B-456; *prtR* proteinase in strains O14, UD2202, UD1001, N1110, UD193, Mo2, while only strain DIALYAC showed the presence of the cell wall-associated high-molecular-mass polysaccharide biosynthesis *cps1 A-J* found in *Lcb. casei* Shirota, and strain DSM4905 xylose isomerase *xylA* gene.

Table 5, probiotic genes presence analysis, the green coloured boxes indicate the presence in the genome



Plasmids

The presence of plasmidic sequences was detected in 7 strains with a total length ranging between 3190 and 16389 bp, with a number of CDS between 3 and 26 (**Table 6**). Analyzing the CDS with the help of Blast, the identified sequences were found to be almost all replication proteins, replication initiators, and relaxases. ATPases, metallophosphoesterase, CPBP intramembrane metalloprotease, membrane mannoprotein, transposase, crystallin protein, Yxea protein, restriction endonuclease S, SMI1/KNR4, peptide cleavage/export ABC transporter, and a protein for bacteriocin secretion have also been predicted. In addition, the presence of PASTA domain-containing protein in strain 54 should be underlined.

Prophages and Crispr-CAS

While the use of Prophage hunter did not show the presence of any prophage, PHASTER highlighted the presence of 7 intact prophages within 5 genomes, and the presence of 26 questionable phages in 10 genomes (**Table 7**). The most frequently detected prophage is PL1.

6 CRISPR-CAS sequences were identified in 4 genomes, mainly CAS-Type IIA systems, together with an IE and an IC therefore both type I and type II (**Table 8**).

Table 6, plasmid predicted by Recycler

Strain	RNODE	bp	total bp	CDS	Predicted proteins
DSM4905	1	7413	7413	5	metallophosphoesterase
NRRL B-456	2 3 4 5	5087 7017 4982 6184	23270	26	cell wall mannoprotein, transposase, Hsp20/alpha crystallin family protein, YxeA, PASTA domain-containing, restriction endonuclease S, SMI1/KNR4
M268	1	11911	11911	12	CPBP intermembrane metalloprotease, peptide cleavage/export ABC transporter, bacteriocin secretion protein,
I2	1 2 3 5	2811 3148 5495 4935	16389	14	
TMW 1.300	1	3190	3190	4	
N1110	1 4	8902 3149	12051	11	LCAM36_0075/DUF1906 domain, membrane protein, DUF536,
N202	2	3190	3190	3	

Table 7, prophages predicted with PHASTER

	Lenght (Kb)	Completeness	Phage	
LMG 25883	31,5	Intact	BH1	NC_005893
	51,1	Intact	PL 1	NC_031125
	69,3	Questionable	iLp84	NC_022757
DSM4905	9	Intact	PBL1c	NC_048689
	49,1	Questionable	CL1	NC_028888
	44,3	Questionable	PLE3	NC_031125
NRRL B-456	14,2	Questionable	LJ	NC_048680
M268	42,9	Intact	PL 1	NC_022757
	34	Intact	iA2	NC_028830
O14	16,4	Intact	SPbeta_like	NC_029119
	37,6	Questionable	Lrm1	NC_011104
UD2202	47	Intact	BH1	NC_048737
	65,9	Questionable	PLE3	NC_031125
	16,1	Questionable	Strept_315.2	NC_004585
I2	34,5	Questionable	PLE3	NC_031125
	6,4	Questionable	RCS47	NC_042128
UD1001	13	Intact	phiAT3	NC_005893
	15,5	Questionable	IME_EFm5	NC_028826
N1110	51,9	Intact	Lrm1	NC_011104
	90,2	Questionable	PLE3	NC_031125
N202	43,8	Intact	PL 1	NC_022757
UD193	40,6	Questionable	iLp84	NC_028783
	15,3	Questionable	SPbeta_like	NC_029119
Mo2	41	Questionable	PLE3	NC_031125
	32,1	Questionable	T25	NC_048625
TMW 1.300	42,3	Intact	BH1	NC_048737
	24,7	Questionable	SPbeta_like	NC_029119
	55	Questionable	PLE3	NC_031125
DIALYac	19	Intact	iA2	NC_028830
	24,7	Intact	PLE2	NC_031036
	36,7	Intact	PLE2	NC_031036
	13,6	Questionable	IME_EFm5	NC_028826

Table 8, CRISPR/Cas predicted with CRISPRCasFinder

Strain	ID	Spacer/ Gene	Repeat consensus/cas genes	Evidence Level
LMG 25883	CRISPR	38	gctcttgaactgattgattcgacatctacctgagac	4
	CAS-Type IIA	4	csn2_TypeIIA, cas2_TypeI-II-III, cas1_TypeII, cas9_TypeII	
NRRL B-456	CRISPR	44	gctcttgaactgattgattcgacatctacctgagac	4
	CAS-Type IIA	4	csn2_TypeIIA, cas2_TypeI-II-III, cas1_TypeII, cas9_TypeII	
	CRISPR	80	gttttccccgcacatgcgggggtgatcc	4
	CAS-Type IE	7	cas3_TypeI, cse2_TypeIE, cas7_TypeIE, cas5_TypeIE, cas6_TypeIE, cas1_TypeIE, cas2_TypeIE	
O14	CRISPR	12	gttcttgaactgattgatctgacatctacctgagac	4
	CAS-Type IIA	4	csn2_TypeIIA, cas2_TypeI-II-III, cas1_TypeII, cas9_TypeII	
UD2202	CRISPR	12	atttcaattcacgcagtcacgtagactgcgac	4
	CRISPR	11	gtcgcagtccacgtgactgcgtgaattgaaat	4
	CAS Type IC	7	cas4_TypeI-II, cas1_TypeIC, cas2_TypeI-II-III, cas5c_TypeIC, cas8c_TypeIC, cas7c_TypeIC, cas3_TypeI	
	CRISPR	26	gctcttgaactgattgatctgacatctacctgagac	4
	CAS Type IIA	4	csn2_TypeIIA, cas2_TypeI-II-III, cas1_TypeII, cas9_TypeII	

Bacteriocins

Through BAGEL4, 49 database hits of presumed bacteriocin were obtained, none of which, however, with an exact match. The most predicted sequences in the genomes were carnocin CP52 and enterocin x chain beta (**Table 9)**.

Table 9, Bacteriocins predicted with BAGEL4

Prediction	Strains
Acidocin_LF221B(GassericinK7B)	UD1001
Carnocin_CP52	LMG 25883, M268, O14, UD2202, I2, UD1001, N1110, N202, UD193, Mo2, TMW 1.300, DIALYac
Enterocin_1071B	UD2202
Enterocin_X_chain_beta	LMG 25883,DSM4905, NRRL B-456, M268, O14, UD2202, I2, UD1001, N1110, N202, UD193, Mo2, TMW 1.300, DIALYac
LSEI_2386	O14, N1110, N202, UD193, Mo2
Sakacin_P_(Sakacin674)	UD2202, UD1001
Salivaricin_A3	UD2202
Thermophilin_A	TMW 1.300

5.5 Discussion and conclusions

In this work, 14 genomes previously identified by other techniques were analysed *in silico* for the entire genomic sequence to verify the correspondence between the different identification methods (Iacumin et al., 2015). Using dDDH and ANI, all the strains clustered with the correspondent reference strains, making the identification clear accordingly to these adopted parameters (21); (46);(47);(48); (49). In contrast, the previous work showed an ambiguous classification for strain DSM4905 between *Lcb. casei* and *Lcb. paracasei*, and for strain UD2202 that was classified as *Lcb. casei* instead of *Lcb. zeae*. With the current approach, strain DSM4905 clustered very tightly within *Lcb. paracasei*, and in proximity with *Lcb. paracasei* subsp. *tolerans* (previously *L. casei* subsp. *tolerans*). This was probably the reason for the doubtful previous identification for strain DSM4905, even if such problematic was not observed for strain M268, also very close to *Lcb. paracasei* subsp. *tolerans*. As a result, the comparison of ANI and dDDH indicates the effectiveness of whole-genome analysis in providing additional clues to classification and in resolving doubts about unclear results of other molecular techniques, with the advantage of considering a larger number of references to be compared for characteristics.

The classification of strains UD2202 and UD1001 as *Lcb. zeae* and *Lcb. casei,* respectively*,* gave less marked taxonomic parameters. This is due to the closeness between these two species, which in

fact were merged (5), and recently re-separated (4). From the analysis of the genomes in **Figure 3** these two species possess a high similarity presenting a core genome with a high number of common genes. Given that the attribution to one species or another in this case is based on a very narrow range of values (1 point of ANI), it is critical that future studies focus on the various genes that characterize these species in order to reconstruct details for which genetic and metabolic traits *Lcb. casei* and *Lcb. zeae* differ from each other, thus providing additional clues for a correct taxonomic attribution. However, for this purpose it is essential to reanalyze and correctly re-identify the strains available in literature, in order not to compromise the future results. Thus, as described by Kim et al. (2021**),** the presence of glycosyltransferase family 8 (KRK10099.1) considered characteristic and uniquely present in the genome of strains belonging to *Lcb. zeae*, was found also in other strains belonging to the *Lcb. casei* species. This erroneous conclusion could be attributed to the taxonomic identification errors that distinguishes the strains deposited for these species as evidenced by **Figure 3**.

Concerning the genetic characterization, neither the alignment on specific databases nor the alignment on blasts of the sequences reported in **Table 1** revealed the presence of potential virulence genes. However, from the annotation of Prokka, the presence of efflux pump systems with drug resistance function NorB in the pan-genome of *Lcb. rhamnosus* and *Lcb. casei*, and AbcA in *Lcb. paracasei*, were predicted. Before being able to consider these strains completely safe, the real presence and expression of these efflux pumps, and their effective efficacy on drug resistance, must therefore be verified. Furthermore, always with the help of Prokka for the identification of the ORF, and blast for the correct identification, a PASTA domain-containing protein was identified (Yeats et al., 2002). This protein in some species was associated with resistance to antibiotics and other cell wall stress responses, as reported in the case of enterococcal colonization of the mammalian gut (Djorić et al., 2021) (Pensinger et al., 2018). The fact that this factor associated with episodes of resistance has not been identified by the antibiotic-resistance detection tools, but only by manual search via blast, underlines the fact that the search for potential virulence factors through these databases can only provide a clue on the possible absence of certain factors, but not from a total certainty on the safety of microorganisms. The identification of potential virulence and resistance genes is critical because the use of these bacteria as probiotics, particularly in immunocompromised people, could result in the onset of infections and adverse phenomena, as has been reported in several cases (34). Further consideration should be given to the possibility of transferring these factors to other bacteria, as was discovered in our case with the presence on plasmids, in order to prevent the spread and transmission of these factors to other bacteria, which could potentially become more pathogenic and virulent. The presence of type II-A, I-E, and I-C CRISPR systems was consistent with what has been reported in

the literature (50) and underlines the ability of these bacteria to develop defense mechanisms against phages, the presence of which as prophages has also been predicted in the 14 genomes studied.

As regards the inhibition capacity of other bacterial species, no known bacteriocins have been identified with certainty, but some possible peptides of interest have been predicted. However, this information is only a preliminary clue, given the low bit score that does not allow a certain identification of the peptide. BAGEL4 aims to suggest the greatest number of positions that can potentially correspond to a bacteriocin, with the possibility of creating false positives. Therefore, the real accuracy of the aforementioned predictions and the completeness of the operon necessary for proper functioning and regulation requires further verifications. Finally, the WGS technique has allowed for the clarification of identification aspects that previously cast doubt on the strains' specific identification. Unfortunately, however, there are still some problems that depend on the incorrect nomenclature of the deposited strains. Therefore it becomes of fundamental importance to proceed with an immediate correction of errors, finding a system that allows the free deposit of the strains genomes, but which guarantees that new errors are not protracted or inserted. This, will be increasingly important in the future to ensure correctness of scientific results, not only for the identification of new strains or the deepening of their genetic study, but also for metagenomics and studies of populations in diverse environments.

5.6 References

1. Hansen PA, Lessel EF. 1971. Lactobacillus casei (Orla- Jensen) comb. nov. Int J syste 21:69–71.

2. Dicks LMT, Plessis EM Du, Dellagl F, Lauer E. 1996. norn . rev ., Designation of ATCC 334 as the Neotype. Int J Syst Evol Microbiol 46:337–340.

3. Dellaglio F, Felis GE, Torriani S. 2002. The status of the species Lactobacillus casei (Orla-Jensen 1916) Hansen and lessel 1971 and Lactobacillus paracasei Collins et al. 1989. Request for an opinion. Int J Syst Evol Microbiol 52:285–287.

4. Liu DD, Gu CT. 2020. Proposal to reclassify lactobacillus zhaodongensis, lactobacillus zeae, lactobacillus argentoratensis and lactobacillus buchneri subsp. Silagei as lacticaseibacillus zhaodongensis comb. nov., lacticaseibacillus zeae comb. nov., lactiplantibacillus argento. Int J Syst Evol Microbiol 70:6414–6417.

5. Zheng J, Wittouck S, Salvetti E, Franz CMAP, Harris HMB, Mattarelli P, O'toole PW, Pot B, Vandamme P, Walter J, Watanabe K, Wuyts S, Felis GE, Gänzle MG, Lebeer S. 2020. 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 70:2782–2858.

6. Tynkkynen S, Satokari R, Saarela M, Mattila-Sandholm T, Saxelin M. 2001. Comparison of ribotyping, randomly amplified polymorphic DNA analysis, and pulsed-field gel electrophoresis in typing of Lactobacillus rhamnosus and L. casei Strains. VTT Publ 65:69–75.

7. Vásquez A, Molin G, Pettersson B, Antonsson M, Ahrné S. 2005. DNA-based classification and sequence heterogeneities in the 16S rRNA genes of Lactobacillus casei/paracasei and related species. Syst Appl Microbiol 28:430–441.

8. Huang CH, Li SW, Huang L, Watanabe K. 2018. Identification and classification for the Lactobacillus casei group. Front Microbiol 9:1–13.

9. Huang CH, Huang L. 2018. Rapid species- and subspecies-specific level classification and identification of Lactobacillus casei group members using MALDI Biotyper combined with ClinProTools. J Dairy Sci 101:979–991.

10. Bottari B, Felis GE, Salvetti E, Castioni A, Campedelli I, Torriani S, Bernini V, Gatti M. 2017. Effective identification of Lactobacillus casei group species: Genome-based selection of the gene mutL as the target of a novel multiplex PCR assay. Microbiol (United Kingdom) 163:950–960.

11. Iacumin L, Ginaldi F, Manzano M, Anastasi V, Reale A, Zotta T, Rossi F, Coppola R, Comi G. 2015. High resolution melting analysis (HRM) as a new tool for the identification of species belonging to the Lactobacillus casei group andcomparison with species-specific PCRs and multiplex PCR. Food Microbiol 46:357–367.

12. Huang CH, Lee FL. 2011. The dnaK gene as a molecular marker for the classification and discrimination of the Lactobacillus casei group. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol 99:319–327.

13. Huang CH, Chang MT, Huang L, Chu WS. 2015. The dnaJ gene as a molecular discriminator to differentiate among species and strain within the Lactobacillus casei group. Mol Cell Probes 29:479–484.

14. Huang CH, Chang MT, Huang L. 2014. Use of highly variable gene (yycH) as DNA marker to resolve interspecific relationships within the Lactobacillus casei group and a target for developing novel species-specific PCR primers. Eur Food Res Technol 239:719–724.

15. Toh H, Oshima K, Nakano A, Takahata M, Murakami M, Takaki T, Nishiyama H, Igimi S, Hattori M, Morita H. 2013. Genomic Adaptation of the Lactobacillus casei Group. PLoS One 8.

16. Wuyts S, Wittouck S, De Boeck I, Allonsius CN, Pasolli E, Segata N, Lebeer S. 2017. Large-Scale Phylogenomics of the Lactobacillus casei Group Highlights Taxonomic Inconsistencies and Reveals Novel Clade-Associated Features . mSystems 2.

17. FAO/WHO. 2002. Guidelines for the Evaluation of Probiotics in Food 1–11.

18. EFSA. 2021. EFSA statement on the requirements for whole genome sequence analysis of microorganisms intentionally used in the food chain. EFSA J 19.

19. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS, Rooney AP, Yi H, Xu XW, De Meyer S, Trujillo ME. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int J Syst Evol Microbiol 68:461–466.

20. Auch AF, Klenk HP, Göker M. 2010. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. Stand Genomic Sci 2:142–148.

21. Auch AF, von Jan M, Klenk HP, Göker M. 2010. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2:117–134.

22. Steinegger M, Salzberg SL. 2020. Terminating contamination: Large-scale search identifies more than 2,000,000 contaminated entries in GenBank. Genome Biol 21:1–12.

23. Rossi N, Colautti A, Iacumin L, Piazza C. 2021. WGA-LP: a pipeline for whole genome assembly of contaminated reads. Bioinformatics 1–3.

24. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477.

25. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.

26. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055.

27. Rozov R, Kav AB, Bogumil D, Shterzer N, Halperin E, Mizrahi I, Shamir R. 2017. Recycler: An algorithm for detecting plasmids from de novo assembly graphs. Bioinformatics 33:475–482.

28. Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun 10.

29. Rodriguez-R L, Konstantinidis K. 2016. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes https://doi.org/10.7287/peerj.preprints.1900.

30. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32:929–931.

31. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. 2006. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res 34:32–36.

32. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen AL V., Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran HK, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar G V., McArthur AG. 2020. CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 48:D517–D525.

33. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, Philippon A, Allesoe RL, Rebelo AR, Florensa AF, Fagelhauer L, Chakraborty T, Neumann B, Werner G, Bender JK, Stingl K, Nguyen M, Coppens J, Xavier BB, Malhotra-Kumar S, Westh H, Pinholt M, Anjum MF, Duggett NA, Kempf I, Nykäsenoja S, Olkkola S, Wieczorek K, Amaro A, Clemente L, Mossong J, Losch S, Ragimbeau C, Lund O, Aarestrup FM. 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother 75:3491–3500.

34. Colautti A, Arnoldi M, Comi G, Iacumin L. 2021. Antibiotic resistance and virulence factors in lactobacilli: Something to carefully consider. Food Microbiol 103:103934.

35. Lebeer S, Vanderleyden J, De Keersmaecker SCJ. 2008. Genes and Molecules of Lactobacilli Supporting Probiotic Action. Microbiol Mol Biol Rev 72:728–764.

36. Goel A, Halami PM, Tamang JP. 2020. Genome Analysis of Lactobacillus plantarum Isolated From Some Indian Fermented Foods for Bacteriocin Production and Probiotic Marker Genes. Front Microbiol 11:1–12.

37. Papadimitriou K, Zoumpopoulou G, Foligné B, Alexandraki V, Kazou M, Pot B, Tsakalidou E. 2015. Discovering probiotic microorganisms: In vitro, in vivo, genetic and omics approaches. Front Microbiol 6:1–28.

38. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–W21.

39. Song W, Sun HX, Zhang C, Cheng L, Peng Y, Deng Z, Wang D, Wang Y, Hu M, Liu W, Yang H, Shen Y, Li J, You L, Xiao M. 2019. Prophage Hunter: an integrative hunting tool for active prophages. Nucleic Acids Res 47:W74–W80.

40. Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res 35:52–57.

41. Seemann T. 2014. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.

42. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693.

43. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Huerta-Cepas J. 2021. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. bioRxiv 5:2021.06.03.446934.

44. Kanehisa M, Sato Y. 2020. KEGG Mapper for inferring cellular functions from protein sequences. Protein Sci 29:28–35.

45. Kim E, Yang S, Kim H. 2021. Differentiation of Lacticaseibacillus zeae Using Pan-Genome Analysis and Real-Time PCR Method Targeting a Unique Gene.

46. Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A 106:19126–19131.

47. Ciufo S, Kannan S, Sharma S, Badretdin A, Clark K, Turner S, Brover S, Schoch CL, Kimchi A, DiCuccio M. 2018. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. Int J Syst Evol Microbiol 68:2386–2392.

48. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:1–8.

49. Wang Q, Zhu W, Wang E, Zhang L, Li X, Wang G. 2016. Genomic identification of rhizobia-related strains and threshold of ANI and core-genome for family, genus and species. Int J Environ Agric Res ISSN 2:76–86.

50. Pujato S, Galliani V, Irazoqui JM, Amadío A, Quiberoni A, Mercanti D. 2021. Analysis of CRISPR systems of types II-A, I-E and I-C in strains of Lacticaseibacillus. Int Dairy J 118.

5.7 Supplementary material

Supplementary Table S1, NCBI accession numbers for the DNA sequence of the strains employed in this work

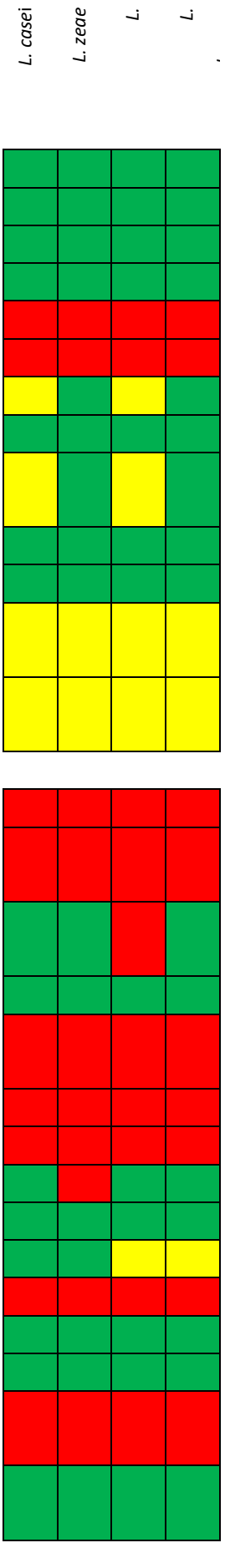
L. casei	*L. paracasei*			*L. rhamnosus*	
NZ_CP006690.1	NZ_CP031785.1	NZ_AP018392.1	NZ_CP041944.1	NZ_CP021426.1	NZ_CP067365.1
NZ_AP012544.1	NZ_CP052065.1	NZ_CP014985.1	NZ_CP068408.1	NZ_CP040780.1	NC_013199.1
CP074377.1	NC_022112.1	NZ_AP012541.1	NZ_CP068416.1	NC_017491.1	NC_021723.1
NZ_CP017065.1	NZ_CP025499.1	NZ_CP013921.1	NZ_CP035563.1	NZ_CP067042.1	NC_021725.1
NZ_CP064303.1	NC_008526.1	NZ_CP012148.1	NZ_CP038153.1	NZ_CP014201.1	NZ_CP017063.1
	NC_017474.1	NC_017473.1	NZ_CP044361.1	NZ_CP046267.1	CP016823.1
	NC_010999.1	NZ_CP029536.1	NZ_CP017716.1	NZ_CP044506.1	NZ_CP025428.1
	NZ_CP048003.1	NC_021721.1	NZ_CP045567.1	NZ_LT220504.1	NZ_CP053619.1
	NZ_CP012187.1	NZ_CP039707.1	NZ_CP016355.1	NZ_CP073317.1	NZ_LR698954.1
	NZ_CP041657.1	NZ_CP029686.1	NC_018641.1	NZ_CP006804.1	NZ_LR134322.1
	NZ_CP029546.1	NZ_CP064299.1	NZ_CP032637.1	NZ_CP031290.1	NZ_LR134331.1
	NZ_CP017261.1	NZ_CP064311.1	NC_014334.2	NC_017482.1	NZ_CP020464.1
	NZ_CP025582.1	NZ_CP064304.1	NZ_CP065154.1	NC_013198.1	NZ_CP019305.1
	NZ_CP026097.1	NZ_CP007122.1	NZ_CP064314.1	NZ_CP046395.1	NZ_CP045586.1
	NZ_CP072181.1	NZ_LR698988.1	NZ_CP064316.1	NZ_CP044228.1	NZ_CP073711.1
	NZ_CP022954.1	NZ_CP050500.1		NZ_CP022109.1	

Supplementary Table 2, glycosyltransferase KRK10099.1 in *L. casei* and *L. zeae*

Strain	Accession number	Score	Expect	Identities	Positives	Gaps	Frame
FBL6	NZ_CP074377.1	613 bits(1581)	0.0	294/299(98%)	298/299(99%)	0/299(0%)	2
LC5	NZ_CP017065.1	182 bits(463)	8E-53	105/303(35%)	150/303(49%)	33/303(10%)	1
MGB0470	NZ_CP064303.1	81 bits(458)	4E-52	103/297(35%)	148/297(49%)	21/297(7%)	1
N	NZ_CP077759.1	613 bits(1581)	0.0	294/299(98%)	298/299(99%)	0/299(0%)	1
N87	NZ_LCUN01000013.1	X					
BIO5773	NZ_WBOC01000001.1	X					
L.cR4	NZ_JAAQWB010000001.1	X					
GCRL 163	NZ_MODT01000077.1	X					
MJA 12	NZ_MODS01000096.1	X					
DS1_13	NZ_QAZD01000001.1	X					
HUL 5	NZ_JAGDFA010000001.1	x					
HUL 12	NZ_JAGEPP010000001.1	x					
DS13_13	NZ_QAZE01000001.1	X					
UBLC-42	NZ_JADPYW010000001.1	X					
Z11	NZ_MPOP01000001.1	X					
UW1	NZ_JDWK01000001.1	x					
B900021	NZ_LOJN01000184.1	x					
AMBR2	FXZN01000001.1	x					
NBRC 101979	NZ_BJUH01000001.1	x					
FAM 20446	NZ_VBSQ01000001.1	x					
867_LCAS	NZ_JUPZ01000059.1	x					
YNF-5	NZ_SDJZ01000001.1	x					
BCRC 80156	NZ_VBWM01000001.1	x					

BCRC 17487	NZ_VBWL01000001.1	x					
12A	NZ_CP006690.1	56.2 bits(134)	1E-09	53/232(23%)	101/232(43%)	14/232(6%)	3
21/1	AFYK01000027.1	x					
A2-362	NZ_AZOE01000001.1	x					
ATCC 393	NZ_AP012544.1	181 bits(458)	4E-52	103/297(35%)	148/297(49%)	21/297(7%)	2
JCM 1134	BALS01000001.1	x					
LcA	CM001861.1	58.9 bits(141)	1E-10	54/232(23%)	102/232(43%)	14/232(6%)	1
UW4	NZ_AFYS01000048.1	x					
ZEAE FBL8	NZ_CP074379.1	621 bits(1602)	0.0	298/299(99%)	299/299(100%)	0/299(0%)	-3
MGYG-HGUT-02383	NZ_CABMJL010000013.1	x					
CRBIP24.58	NZ_VBWN01000001.1	621 bits(1602)	0.0	298/299(99%)	299/299(100%)	0/299(0%)	-3
CRBIP24.44	NZ_VBWO01000001.1	x					
CECT 9104	NZ_LS991421.1	613 bits(1581)	0.0	294/299(98%)	297/299(99%)	0/299(0%)	-1
KCTC 3804	NZ_BACQ01000031.1	X					
DSM 20178	NZ_AZCT01000001.1	X					

Supplementary material Table S3, metabolisms analysis made with KEGG



Carbohydrate metabolism

Central carbohydrate metabolism

M00001 Glycolysis (Embden-Meyerhof pathway), glucose => pyruvate (11) (complete 9/9)

M00002 Glycolysis, core module involving three-carbon compounds (6) (complete 5/5)

M00003 Gluconeogenesis, oxaloacetate => fructose-6P (9) (complete 7/7)

M00307 Pyruvate oxidation, pyruvate => acetyl-CoA (4) (complete 1/1)

M00009 Citrate cycle (TCA cycle, Krebs cycle) (4) (incomplete 1/8)

M00011 Citrate cycle, second carbon oxidation, 2-oxoglutarate => oxaloacetate (4) (incomplete 1/5)

M00004 Pentose phosphate pathway (Pentose phosphate cycle) (7) (1 block missing 6/7)

M00006 Pentose phosphate pathway, oxidative phase, glucose 6P => ribulose 5P (3) (complete 2/2)

M00007 Pentose phosphate pathway, non-oxidative phase, fructose 6P => ribose 5P (3) (1 block missing 3/4)

M00580 Pentose phosphate pathway, archaea, fructose 6P => ribose 5P (3) (complete 2/2)

M00005 PRPP biosynthesis, ribose 5P => PRPP (1) (complete 1/1)

M00008 Entner-Doudoroff pathway, glucose-6P => glyceraldehyde-3P + pyruvate (3) (1 block missing 3/4)

M00308 Semi-phosphorylative Entner-Doudoroff pathway, gluconate => glycerate-3P (4) (1 block missing 3/4)

Other carbohydrate metabolism

M00014 Glucuronate pathway (uronate pathway) (3) (incomplete 2/7)

M00631 D-Galacturonate degradation (bacteria), D-galacturonate => pyruvate + D-glyceraldehyde 3P (4) (2 blocks missing 3/5)

M00061 D-Glucuronate degradation, D-glucuronate => pyruvate + D-glyceraldehyde 3P (7) (complete 5/5)

M00632 Galactose degradation, Leloir pathway, galactose => alpha-D-glucose-1P (4) (complete 4/4)

M00552 D-galactonate degradation, De Ley-Doudoroff pathway, D-galactonate => glycerate-3P (3) (2 blocks missing 3/5)

M00129 Ascorbate biosynthesis, animals, glucose-1P => ascorbate (2) (incomplete 2/7)

M00114 Ascorbate biosynthesis, plants, fructose-6P => ascorbate (1) (incomplete 1/8)

M00550 Ascorbate degradation, ascorbate => D-xylulose-5P (7) (complete 5/5)

M00854 Glycogen biosynthesis, glucose-1P => glycogen/starch (4) (complete 2/2)

M00855 Glycogen degradation, glycogen => glucose-6P (3) (1 block missing 2/3)

M00565 Trehalose biosynthesis, D-glucose 1P => trehalose (3) (incomplete 3/6)

M00549 Nucleotide sugar biosynthesis, glucose => UDP-glucose (3) (complete 3/3)

M00554 Nucleotide sugar biosynthesis, galactose => UDP-galactose (2) (complete 2/2)

M00892 UDP-N-acetyl-D-glucosamine biosynthesis, eukaryotes, glucose => UDP-GlcNAc (3) (incomplete 3/6)

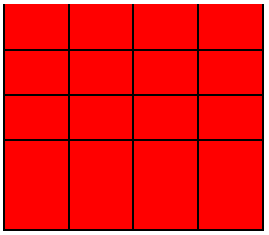
M00909 UDP-N-acetyl-D-glucosamine biosynthesis, prokaryotes, glucose => UDP-GlcNAc (5) (complete 5/5)

M00373 Ethylmalonyl pathway (1) (incomplete 1/12)

M00532 Photorespiration (4) (incomplete 2/10)

M00013 Malonate semialdehyde pathway, propanoyl-CoA => acetyl-CoA (1) (incomplete 1/5)

M00131 Inositol phosphate metabolism, Ins(1,3,4,5)P4 => Ins(1,3,4)P3 => myo-inositol (1) (incomplete 1/4)



Energy metabolism

Carbon fixation

M00165 Reductive pentose phosphate cycle (Calvin cycle) (5) (incomplete 7/11)

M00166 Reductive pentose phosphate cycle, ribulose-5P => glyceraldehyde-3P (2) (2 blocks missing 2/4)

M00167 Reductive pentose phosphate cycle, glyceraldehyde-3P => ribulose-5P (3) (2 blocks missing 5/7)

M00169 CAM (Crassulacean acid metabolism), light (1) (1 block missing 1/2)

M00172 C4-dicarboxylic acid cycle, NADP - malic enzyme type (1) (incomplete 1/4)

M00171 C4-dicarboxylic acid cycle, NAD - malic enzyme type (1) (incomplete 1/7)

M00170 C4-dicarboxylic acid cycle, phosphoenolpyruvate carboxykinase type (1) (incomplete 1/4)

M00173 Reductive citrate cycle (Arnon-Buchanan cycle) (5) (incomplete 2/10)

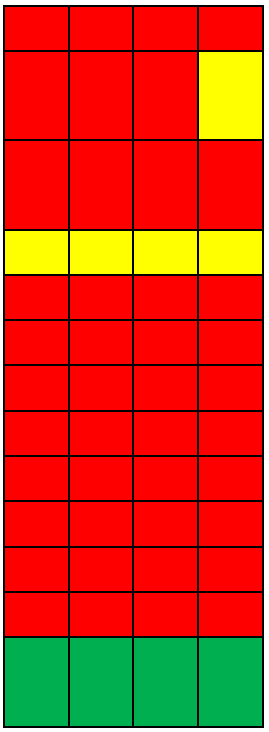
M00376 3-Hydroxypropionate bi-cycle (6) (incomplete 2/13)

M00375 Hydroxypropionate-hydroxybutylate cycle (1) (incomplete 1/14)

M00374 Dicarboxylate-hydroxybutyrate cycle (2) (incomplete 1/13)

M00377 Reductive acetyl-CoA pathway (Wood-Ljungdahl pathway) (3) (incomplete 3/7)

M00579 Phosphate acetyltransferase-acetate kinase pathway, acetyl-CoA => acetate (2) (complete 2/2)



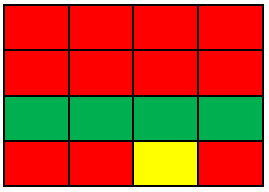
Methane metabolism

M00357 Methanogenesis, acetate => methane (2) (incomplete 1/5)

M00346 Formaldehyde assimilation, serine pathway (3) (incomplete 3/9)

M00345 Formaldehyde assimilation, ribulose monophosphate pathway (4) (complete 3/3)

M00344 Formaldehyde assimilation, xylulose monophosphate pathway (1) (incomplete 1/4)



ATP synthesis

M00157 F-type ATPase, prokaryotes and chloroplasts (8) (complete 1/1)

Lipid metabolism

Fatty acid metabolism

M00082 Fatty acid biosynthesis, initiation (6) (complete 2/2)

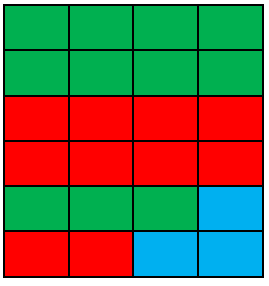
M00083 Fatty acid biosynthesis, elongation (4) (complete 1/1)

M00873 Fatty acid biosynthesis in mitochondria, animals (2) (incomplete 1/6)

M00874 Fatty acid biosynthesis in mitochondria, fungi (3) (incomplete 2/6)

M00086 beta-Oxidation, acyl-CoA synthesis (1) (complete 1/1)

M00087 beta-Oxidation (1) (2 blocks missing 1/3)

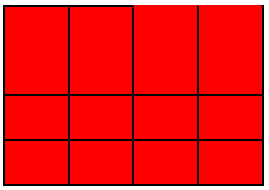


Lipid metabolism

M00088 Ketone body biosynthesis, acetyl-CoA => acetoacetate/3-hydroxybutyrate/acetone (2) (incomplete 2/5)

M00089 Triacylglycerol biosynthesis (1) (incomplete 1/4)

M00093 Phosphatidylethanolamine (PE) biosynthesis, PA => PS => PE (1) (2 blocks missing 1/3)



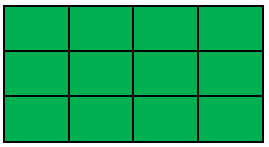
Nucleotide metabolism

Purine metabolism

M00048 Inosine monophosphate biosynthesis, PRPP + glutamine => IMP (10) (complete 8/8)

M00049 Adenine ribonucleotide biosynthesis, IMP => ADP,ATP (5) (complete 4/4)

M00050 Guanine ribonucleotide biosynthesis, IMP => GDP,GTP (5) (complete 4/4)



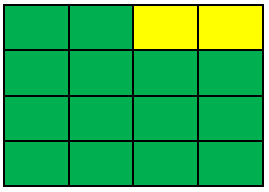
Pyrimidine metabolism

M00051 Uridine monophosphate biosynthesis, glutamine (+ PRPP) => UMP (9) (1 block missing 2/3)

M00052 Pyrimidine ribonucleotide biosynthesis, UMP => UDP/UTP,CDP/CTP (3) (complete 3/3)

M00053 Pyrimidine deoxyribonuleotide biosynthesis, CDP => dCTP (3) (complete 2/2)

M00938 Pyrimidine deoxyribonuleotide biosynthesis, UDP => dTTP (6) (complete 5/5)



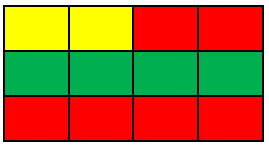
Amino acid metabolism

Serine and threonine metabolism

M00020 Serine biosynthesis, glycerate-3P => serine (1) (2 blocks missing 1/3)

M00018 Threonine biosynthesis, aspartate => homoserine => threonine (6) (complete 5/5)

M00033 Ectoine biosynthesis, aspartate => ectoine (2) (incomplete 2/5)



Cysteine and methionine metabolism

M00021 Cysteine biosynthesis, serine => cysteine (1) (1 block missing 1/2)

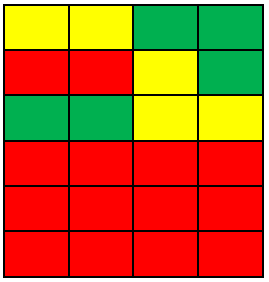
M00609 Cysteine biosynthesis, methionine => cysteine (4) (2 blocks missing 4/6)

M00017 Methionine biosynthesis, apartate => homoserine => methionine (7) (1 block missing 6/7)

M00034 Methionine salvage pathway (4) (incomplete 2/8)

M00035 Methionine degradation (2) (2 blocks missing 2/4)

M00368 Ethylene biosynthesis, methionine => ethylene (1) (2 blocks missing 1/3)

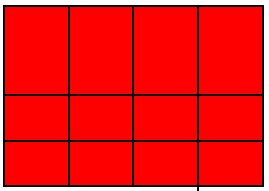


Branched-chain amino acid metabolism

M00019 Valine/isoleucine biosynthesis, pyruvate => valine / 2-oxobutanoate => isoleucine (2) (incomplete 1/4)

M00570 Isoleucine biosynthesis, threonine => 2-oxobutanoate => isoleucine (3) (incomplete 2/5)

M00036 Leucine degradation, leucine => acetoacetate + acetyl-CoA (2) (incomplete 1/6)



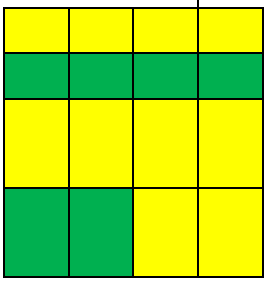
Lysine metabolism

M00016 Lysine biosynthesis, succinyl-DAP pathway, aspartate => lysine (8) (1 block missing 8/9)

M00525 Lysine biosynthesis, acetyl-DAP pathway, aspartate => lysine (9) (complete 9/9)

M00526 Lysine biosynthesis, DAP dehydrogenase pathway, aspartate => lysine (5) (1 block missing 5/6)

M00527 Lysine biosynthesis, DAP aminotransferase pathway, aspartate => lysine (6) (1 block missing 6/7)

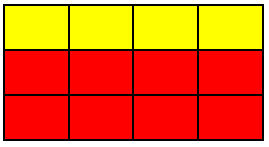


Arginine and proline metabolism

M00844 Arginine biosynthesis, ornithine => arginine (2) (1 block missing 2/3)

M00845 Arginine biosynthesis, glutamate => acetylcitrulline => arginine (2) (incomplete 2/7)

M00029 Urea cycle (2) (incomplete 2/5)

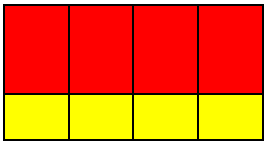


M00015 Proline biosynthesis, glutamate => proline (3) (complete 2/2)

Polyamine biosynthesis

M00133 Polyamine biosynthesis, arginine => agmatine => putrescine => spermidine (1) (incomplete 1/4)

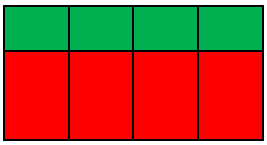
M00134 Polyamine biosynthesis, arginine => ornithine => putrescine (1) (1 block missing 1/2)



Histidine metabolism

M00026 Histidine biosynthesis, PRPP => histidine (12) (complete 6/6)

M00045 Histidine degradation, histidine => N-formiminoglutamate => glutamate (1) (incomplete 1/4)



Aromatic amino acid metabolism

M00023 Tryptophan biosynthesis, chorismate => tryptophan (7) (complete 3/3)

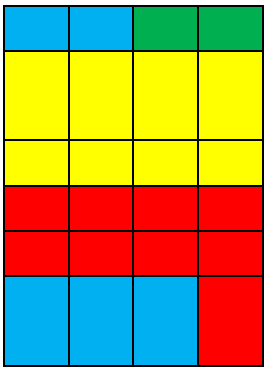
M00024 Phenylalanine biosynthesis, chorismate => phenylpyruvate => phenylalanine (2) (1 block missing 1/2)

M00025 Tyrosine biosynthesis, chorismate => HPP => tyrosine (2) (1 block missing 1/2)

M00040 Tyrosine biosynthesis, chorismate => arogenate => tyrosine (1) (2 blocks missing 1/3)

M00044 Tyrosine degradation, tyrosine => homogentisate (1) (incomplete 1/5)

M00533 Homoprotocatechuate degradation, homoprotocatechuate => 2-oxohept-3-enedioate (1) (incomplete 1/4)



Other amino acid metabolism

M00027 GABA (gamma-Aminobutyrate) shunt (1) (2 blocks missing 1/3)

M00118 Glutathione biosynthesis, glutamate => glutathione (1) (1 block missing 1/2)

Glycan metabolism

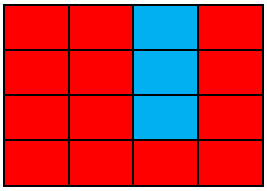
Glycosaminoglycan metabolism

M00076 Dermatan sulfate degradation (1) (incomplete 1/5)

M00077 Chondroitin sulfate degradation (1) (incomplete 1/4)

M00078 Heparan sulfate degradation (1) (incomplete 1/8)

M00079 Keratan sulfate degradation (1) (incomplete 1/4)



Lipopolysaccharide metabolism

M00063 CMP-KDO biosynthesis (1) (incomplete 1/4)

Metabolism of cofactors and vitamins

Cofactor and vitamin metabolism

M00127 Thiamine biosynthesis, prokaryotes, AIR (+ DXP/tyrosine) => TMP/TPP (3) (incomplete 2/7)

M00895 Thiamine biosynthesis, prokaryotes, AIR (+ DXP/glycine) => TMP/TPP (3) (incomplete 3/9)

M00896 Thiamine biosynthesis, archaea, AIR (+ NAD+) => TMP/TPP (2) (incomplete 1/4)

M00897 Thiamine biosynthesis, plants, AIR (+ NAD+) => TMP/thiamine/TPP (1) (incomplete 1/5)

M00898 Thiamine biosynthesis, pyridoxal-5P => TMP/thiamine/TPP (1) (incomplete 1/5)

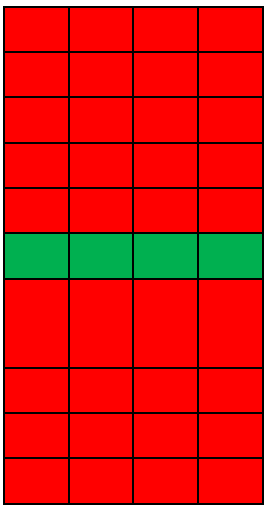
M00899 Thiamine salvage pathway, HMP/HET => TMP (3) (complete 2/2)

M00125 Riboflavin biosynthesis, plants and bacteria, GTP => riboflavin/FMN/FAD (3) (incomplete 2/7)

M00115 NAD biosynthesis, aspartate => quinolinate => NAD (2) (incomplete 2/5)

M00912 NAD biosynthesis, tryptophan => quinolinate => NAD (2) (incomplete 2/8)

M00119 Pantothenate biosynthesis, valine/L-aspartate => pantothenate (2) (incomplete 2/5)



M00913 Pantothenate biosynthesis, 2-oxoisovalerate/spermine => pantothenate (1) (incomplete 1/5)

M00120 Coenzyme A biosynthesis, pantothenate => CoA (5) (complete 3/3)

M00914 Coenzyme A biosynthesis, archaea, 2-oxoisovalerate => 4-phosphopantoate => CoA (2) (incomplete 2/7)

M00572 Pimeloyl-ACP biosynthesis, BioC-BioH pathway, malonyl-ACP => pimeloyl-ACP (3) (incomplete 3/6)

M00883 Lipoic acid biosynthesis, animals and bacteria, octanoyl-ACP => dihydrolipoyl-H => dihydrolipoyl-E2 (1) (2 blocks missing 1/3)

M00126 Tetrahydrofolate biosynthesis, GTP => THF (6) (1 block missing 4/5)

M00840 Tetrahydrofolate biosynthesis, mediated by ribA and trpF, GTP => THF (2) (incomplete 2/6)

M00841 Tetrahydrofolate biosynthesis, mediated by PTPS, GTP => THF (2) (incomplete 2/5)

M00842 Tetrahydrobiopterin biosynthesis, GTP => BH4 (1) (2 blocks missing 1/3)

M00843 L-threo-Tetrahydrobiopterin biosynthesis, GTP => L-threo-BH4 (1) (2 blocks missing 1/3)

M00140 C1-unit interconversion, prokaryotes (3) (complete 3/3)

M00141 C1-unit interconversion, eukaryotes (1) (1 block missing 1/2)

M00868 Heme biosynthesis, animals and fungi, glycine => heme (1) (incomplete 1/8)

M00121 Heme biosynthesis, plants and bacteria, glutamate => heme (1) (incomplete 1/10)

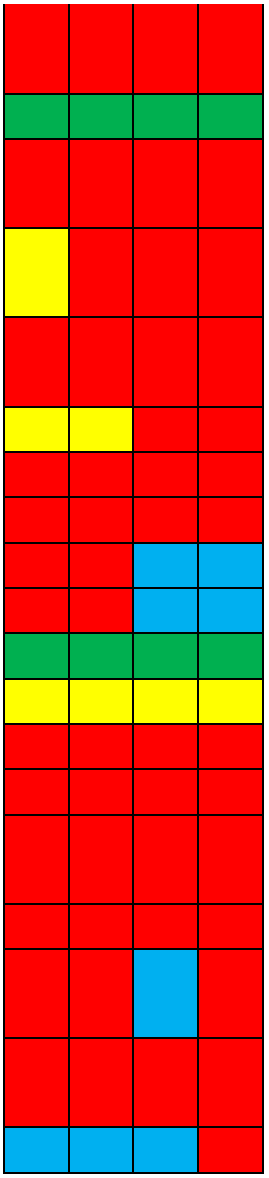
M00926 Heme biosynthesis, bacteria, glutamyl-tRNA => coproporphyrin III => heme (1) (incomplete 1/9)

M00122 Cobalamin biosynthesis, cobyrinate a,c-diamide => cobalamin (1) (incomplete 1/7)

M00117 Ubiquinone biosynthesis, prokaryotes, chorismate (+ polyprenyl-PP) => ubiquinol (2) (incomplete 1/9)

M00116 Menaquinone biosynthesis, chorismate (+ polyprenyl-PP) => menaquinol (1) (incomplete 1/9)

M00932 Phylloquinone biosynthesis, chorismate (+ phytyl-PP) => phylloquinol (3) (incomplete 2/7)



Biosynthesis of terpenoids and polyketides

Terpenoid backbone biosynthesis

M00095 C5 isoprenoid biosynthesis, mevalonate pathway (6) (1 block missing 6/7)

M00849 C5 isoprenoid biosynthesis, mevalonate pathway, archaea (5) (2 blocks missing 4/6)

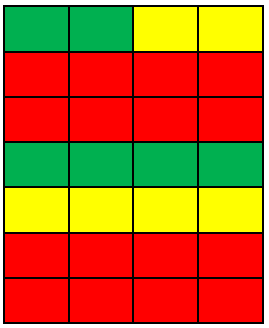
M00096 C5 isoprenoid biosynthesis, non-mevalonate pathway (2) (incomplete 2/8)

M00364 C10-C20 isoprenoid biosynthesis, bacteria (2) (complete 2/2)

M00365 C10-C20 isoprenoid biosynthesis, archaea (1) (1 block missing 1/2)

M00366 C10-C20 isoprenoid biosynthesis, plants (2) (2 blocks missing 2/4)

M00367 C10-C20 isoprenoid biosynthesis, non-plant eukaryotes (1) (2 blocks missing 1/3)



Polyketide sugar unit biosynthesis

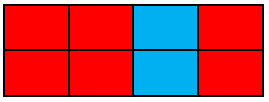
M00793 dTDP-L-rhamnose biosynthesis (1) (2 blocks missing 1/3)

Xenobiotics biodegradation

Aromatics degradation

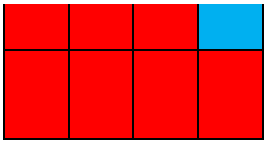
M00538 Toluene degradation, toluene => benzoate (1) (2 blocks missing 1/3)

M00537 Xylene degradation, xylene => methylbenzoate (1) (2 blocks missing 1/3)



M00568 Catechol ortho-cleavage, catechol => 3-oxoadipate (1) (incomplete 1/4)

M00569 Catechol meta-cleavage, catechol => acetyl-CoA / 4-methylcatechol => propanoyl-CoA (2) (incomplete 1/5)



Signature modules

Gene set

Drug resistance

M00627 beta-Lactam resistance, Bla system (2) (1 block missing 2/3)

M00704 Tetracycline resistance, efflux pump Tet38 (1) (1 block missing 1/2)

M00725 Cationic antimicrobial peptide (CAMP) resistance, dltABCD operon (4) (2 blocks missing 1/3)

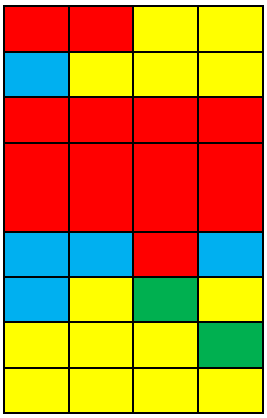
M00726 Cationic antimicrobial peptide (CAMP) resistance, lysyl-phosphatidylglycerol (L-PG) synthase MprF (1) (2 blocks missing 1/3)

M00769 Multidrug resistance, efflux pump MexPQ-OpmE (1) (2 blocks missing 1/3)

M00700 Multidrug resistance, efflux pump AbcA (1) (1 block missing 1/2)

M00702 Multidrug resistance, efflux pump NorB (2) (complete 2/2)

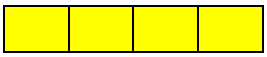
M00705 Multidrug resistance, efflux pump MepA (1) (1 block missing 1/2)



Module set

Metabolic capacity

M00618 Acetogen (0) (1 block missing 1/2



Chapter 6: Draft Genome Sequences of 8 Bacilli isolated from an ancient Roman amphora

Andrea Colautti[a], Giuseppe Comi[a], Emanuele De Paoli[a], Enrico Peterlunger[a], Marta Novello[b], Elena Braidotti[b], Daniele Pasini[b], Lucilla Iacumin[a#]

[a]Department of Agricultural, Food, Environmental and Animal Science (Di4A), University of Udine, 33100 Udine, Italy

[b]Museo archeologico nazionale di Aquileia e Museo paleocristiano di Aquileia, Direzione regionale musei del Friuli Venezia Giulia, via Roma, 1 - 33051 Aquileia (UD)

Running title: Bacilli from ancient Roman amphora

#Corresponding author:

Lucilla Iacumin
Department of Agricultural, Food, Environmental and Animal Science
via Sondrio 2/A, 33100 Udine, Italy,
tel: +390432558126, Fax: +390432558130,
E-mail: lucilla.iacumin@uniud.it

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequences of Eight Bacilli Isolated from an Ancient Roman Amphora

Andrea Colautti,[a] Giuseppe Comi,[a] Emanuele De Paoli,[a] Enrico Peterlunger,[a] Marta Novello,[b] Elena Braidotti,[b] Daniele Pasini,[b] Lucilla Iacumin[a]

[a]Department of Agricultural, Food, Environmental and Animal Science (Di4A), University of Udine, Udine, Italy
[b]Museo archeologico nazionale di Aquileia e Museo paleocristiano di Aquileia, Direzione regionale musei del Friuli-Venezia Giulia, Aquileia, Udine, Italy

ABSTRACT Paleomicrobiology, the study of ancient microbiological material, allows us to understand different evolutionary phenomena in bacteria. In this study, eight bacilli isolated from an ancient Roman amphora, which dates to the IV to V sec. AD, were sequenced and functionally annotated.

Bacteria belonging to *Bacillus* spp. are capable of forming spores, specialized cell forms that can withstand adverse environmental conditions and extreme factors, such as temperature, radiation, and chemicals (1), also allowing for survival in a quiescent state for a long time (2). Isolation of these bacterial species from ancient sources has been previously reported, for example from a mummy (3) or ancient soil (4). In this study, bacilli from an ancient Roman amphora were isolated and sequenced. The amphora (IV to V century AD), found in Aquileia (UD-Italy) (45°45'05.9"N 13°21'03.8"E), was found intact and sealed with cementitious compound, thus preventing microbial contamination. It was opened under aseptic conditions in a laminar flow hood and GMP were followed to avoid contaminations. The inner material was sampled using 10 different culture media for bacteria and fungi by serial dilution method and enrichment steps. Growth (7.26±0.09 log colony forming units/g) was observed only in brain heart infusion and plate count agar (Oxoid, Italy) after 48 h at 30°C under aerobic conditions, showing indented, diffuse mucosal colonies, 1 to 2.5 cm in diameter after 48 h at 30°C. The environmental control made using active/passive methods confirmed the absence of *Bacillus* spp. in the laboratory air. Twenty-five colonies present on the counting plates were isolated and examined for their morphological characteristics, which were Gram- and catalase-positive. Preliminary identification was performed by sequencing amplicons obtained using primers P1 and P4 (5), targeting V1 to V3 regions of 16S rDNA. Amplification conditions: final volume 50 μL, 10 mM Tris–HCl, pH 8, KCl 50 mM, $MgCl_2$ 1.5 mM, dNTPs 0.2 mM, each primer 0.2 μM, 1.25 U *Taq*-polymerase (Applied Biosystem, I), and 100 ng of DNA. After purification, products were sent to a commercial facility for sequencing (Sanger technology, Eurofins Genomics, Germany). Clones were eliminated by comparing genetic fingerprints (by RAPD, Rep-PCR, SAU-PCR) (6) of isolates and the resulting eight unique individual strains were subjected to whole-genome sequencing. For the sequencing process, each strain was cultured in brain heart infusion broth at 30°C for 48 h. After obtaining the cell pellet by centrifugation for 5 min at 5,000 × *g*, the DNA was extracted with the MagAttract HMW DNA Kit (Qiagen, Germany). The DNA was fragmented by sonication (BioRuptor-Diagenode, Belgium) and Celero DNA-Seq kit (Tecan, Swiss) was used for the preparation of libraries. The size of the individual fragments making up the library was measured using BioAnlayzer 2100 DNA chip electrophoresis (Agilent Technologies, USA) and sequencing was carried out with the MiSeq platform (Illumina, USA) in paired-end mode with reads of 300 bp length. The obtained. fastq files were analyzed and assembled using WGA-LP pipeline (7) with the following tools used in default mode. Raw reads were quality trimmed and deprived of Illumina adapters via Trimmomatic v0.39 (8). FastQC v0.11.9 (9) and Kraken2 v2.0.8-b (10)

Editor David Rasko, University of Maryland School of Medicine

Address correspondence to Lucilla Iacumin, lucilla.iacumin@uniud.it.

The authors declare no conflict of interest.

Received 23 March 2022
Accepted 20 May 2022
Published 31 May 2022

6.1 Abstract

Paleomicrobiology, the study of ancient microbiological material, allows us to understand different evolutionary phenomena in bacteria. In this study, eight bacilli isolated from an ancient Roman amphora, which dates to the IV to V sec. AD, were sequenced and functionally annotated.

6.2 Text

Bacteria belonging to Bacillus spp. are capable of forming spores, specialized cell forms that can withstand adverse environmental conditions and extreme factors, such as temperature, radiation, and chemicals (1), also allowing for survival in a quiescent state for a long time (2). Isolation of these bacterial species from ancient sources has been previously reported, for example from a mummy (3) or ancient soil (4). In this study, bacilli from an ancient Roman amphora were isolated and sequenced. The amphora (IV to V century AD), found in Aquileia (UD-Italy) (45°45'05.9" N 13°21'03.8" E), was found intact and sealed with cementitious compound, thus preventing microbial contamination. It was opened under aseptic conditions in a laminar flow hood and GMP were followed to avoid contaminations. The inner material was sampled using 10 different culture media for bacteria and fungi by serial dilution method and enrichment steps. Growth (7.26 ± 0.09 log colony forming units/g) was observed only in brain heart infusion and plate count agar (Oxoid, Italy) after 48 h at 30°C under aerobic conditions, showing indented, diffuse mucosal colonies, 1 to 2.5 cm in diameter after 48 h at 30°C. The environmental control made using active/passive methods confirmed the absence of Bacillus spp. in the laboratory air. Twenty-five colonies present on the counting plates were isolated and examined for their morphological characteristics, which were Gram- and catalase-positive. Preliminary identification was performed by sequencing amplicons obtained using primers P1 and P4 (5), targeting V1 to V3 regions of 16S rDNA. Amplification conditions: final volume 50mL, 10 mM Tris–HCl, pH 8, KCl 50 mM, $MgCl_2$ 1.5mM, dNTPs 0.2mM, each primer 0.2mM, 1.25 U Taq-polymerase (Applied Biosystem, I), and 100 ng of DNA. After purification, products were sent to a commercial facility for sequencing (Sanger technology, Eurofins Genomics, Germany). Clones were eliminated by comparing genetic fingerprints (by RAPD, Rep-PCR, SAU-PCR) (6) of isolates and the resulting eight unique individual strains were subjected to whole-genome sequencing. For the sequencing process, each strain was cultured in brain heart infusion broth at 30°C for 48 h. After obtaining the cell pellet by centrifugation for 5 min at 5,000 × g, the DNA was extracted with the MagAttract HMW DNA Kit (Qiagen, Germany). The DNA was fragmented by sonication (BioRuptor-Diagenode, Belgium) and Celero DNA-Seq kit (Tecan, Swiss) was used for the preparation of libraries. The size of the individual fragments making up the library was measured using BioAnlayzer 2100 DNA chip electrophoresis (Agilent Technologies, USA) and sequencing was

carried out with the MiSeq platform (Illumina, USA) in paired-end mode with reads of 300 bp length. The obtained. fastq files were analyzed and assembled using WGA-LP pipeline (7) with the following tools used in default mode. Raw reads were quality trimmed and deprived of Illumina adapters via Trimmomatic v0.39 (8). FastQC v0.11.9 (9) and Kraken2 v2.0.8-b (10) were used for quality and contamination control. Assembly was carried out using SPAdes v3.15.2 (11). The quality of the final assemblies was evaluated using CheckM v1.1.3 (12), Quast v5.0.2 (13), and SamTools v1.10 (14). Functional annotation was carried out on the genomes using PGAP 2022-04-14.build6021 (15).

Data availability. Sequences were deposited in GenBank with PRJNA811801 BioProject accession number. Table 1 reports the GenBank and SRA accession number, the raw reads number, the NCBI taxonomic identification, the isolation source, the sequencing and assembly statistics, and the genome features of strains for each sample.

Acknowledgments

We thank the Interdepartmental Centre for the Development of the Friulian Language and Culture (CIRF) of the University of Udine for the financial support.

Table 1: Statistics of assembled genomes

GeneBank accession no.	SRA accession no.	16S RNA accession no.	Raw Reads[a]	Strain	Organism name	Source	Coverage	Genome size[b]	Scaffolds[b]	N50[b]	G+C content (%)[c]	CDS[c]	tRNAs[c]	Completeness (%)[d]
JAKXEE000000000	SRR18190504	ON326590	2727696	Aquil_B1	*P. simplex*	Amphora	234×	5649653	25	910271	40.2	5402	81	98,91
JAKXED000000000	SRR18190503	ON326591	717606	Aquil_B2	*L. fusiformis*	Amphora	75×	4643302	34	1011198	37.5	4546	85	99,93
JAKXEC000000000	SRR18190502	ON326592	494295	Aquil_B3	*B. muralis*	Amphora	46×	5057074	38	641321	41.3	4721	84	98,77
JAKXEB000000000	SRR18190501	ON326593	1480745	Aquil_B4	*B. frigoritolerans*	Amphora	65×	6677279	68	317977	39.5	6548	92	98,91
JAKXEA000000000	SRR18190500	ON326594	2409430	Aquil_B5	*B. muralis*	Amphora	151×	5067063	38	641321	41.3	4723	84	98,91
JAKXDZ000000000	SRR18190499	ON326595	1249575	Aquil_B6	*P. psychrodurans*	Amphora	108×	4256356	79	253084	35.9	4213	70	100
JAKXDY000000000	SRR18190498	ON326596	1122132	Aquil_B7	*B. frigoritolerans*	Amphora	71×	5521551	46	613520	40.3	5287	84	98,91
JAKXDX000000000	SRR18190497	ON326597	1742298	Aquil_B8	*P. simplex*	Amphora	57×	5654249	78	193610	40.2	5437	81	98,91

[a] Determined using FastQC

[b] Determined using Quast

[c] Determined using Prokka v1.14.6 with default parameters

[d] Determined using CheckM

6.3 Bibliography

1. Setlow P. 2006. Spores of Bacillus subtilis: Their resistance to and killing by radiation, heat and chemicals. J Appl Microbiol 101:514–525.

2. Setlow P. 2014. Germination of spores of Bacillus species: What we know and do not know. J Bacteriol 196:1297–1305.

3. Čavka M, Glasnović A, Janković I, Šikanjić PR, Perić B, Brkljačić B, Mlinarić-Missoni E, Škrlin J. 2010. Microbiological analysis of a mummy from the archeological museum in Zagreb. Coll Antropol 34:803–805.

4. Southern PM. 2008. Archeology meets clinical microbiology: Analysis of the microbial content of soil from archeological sites in Italy and Belize. Lab Med 39:601–602.

5. Klijn N, Weerkamp AH, De Vos WM. 1991. Identification of mesophilic lactic acid bacteria by using polymerase chain reaction-amplified variable regions of 16S rRNA and specific DNA probes. Appl Environ Microbiol 57:3390–3393.

6. Iacumin L, Comi G, Cantoni C, Cocolin L. 2006. Molecular and technological characterization of Staphylococcus xylosus isolated from naturally fermented Italian sausages by RAPD, Rep-PCR and Sau-PCR analysis. Meat Sci 74:281–288.

7. Rossi N, Colautti A, Iacumin L, Piazza C. 2021. WGA-LP: a pipeline for whole genome assembly of contaminated reads. Bioinformatics 1–3.

8. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

9. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

10. Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. Genome Biol 20:1–13.

11. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477.

12. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055.

13. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.

14. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.

15. Zhao Y, Wu J, Yang J, Sun S, Xiao J, Yu J. 2012. PGAP: Pan-genomes analysis pipeline. Bioinformatics 28:416–418.

Chapter 7: Ancient Roman bacterium against current issues: strain Aquil_B6, *Paenisporosarcina quisquiliarum* or *Psychrobacillus psychrodurans*?

Andrea Colautti[1], Giuseppe Comi[1], Lucilla Iacumin[1]*

[1]Department of Agricultural, Food, Environmental and Animal Science, University of Udine, Udine, Italy

Running Head: Ancient Roman bacteria taxonomic identification

*Corresponding author:
Lucilla Iacumin,
Department of Food Science,
via Sondrio 2/A, 33100 Udine, Italy,
tel: +390432558126, Fax: +390432558130,
E-mail: lucilla.iacumin@uniud.it

7.1 Abstract

Thanks to the development of whole genome sequencing technologies, nowadays researchers have access to rapid, effective, and precise resources for the genetic analysis of prokaryotes. The use of metrics that consider the entire DNA sequence, such as ANI and dDDH calculation, in fact allow better results when compared to single target genes analysis. However, it is increasingly important to employ reference databases containing correct information with high quality sequences. The aim of this work was to taxonomically identify an unknown bacterial strain isolated from an ancient Roman amphora. However, it was necessary to re-sequence the genome of the reference strains Paenisporosarcina quisquiliarum SK 55 and Psychrobacillus psychrodurans DSM 11713 to obtain correct and more reliable sequences than those available. It was therefore possible to uniquely identify the strain Aquil_B6 as a Psychrobacillus psychrodurans. Also, to obtain further genetic information, given the low number of sequences available for this species, the genome of strain DSM 30747 was also sequenced. These new genomes were made publicly available under accession numbers JAMKBI000000000, JAMKBJ000000000 and JAMKBK000000000. A further characterization was made to identify the presence of possible genetic characteristic features such as CRISPR-Cas, prophages, resistance factors, and bacteriocins of this species, making a comparison between the available genomes of Paenisporosarcina quisquiliarum SK 55 and Psychrobacillus psychrodurans strains DSM 11713, DSM 30747, and Aquil_B6, to evaluate the impact of the time in evolution of the considered species.

Importance

Since 1988, through the United States government's founding, the National Center for Biotechnology Information (NCBI) has provided an invaluable service to scientific advancement. The universality and total freedom of use, if on the one hand allow the use of this database on a global level by all researchers for their valuable work, on the other hand it has the disadvantage of making it difficult to check the correctness of all the material present. It is therefore of fundamental importance for the correctness and ethics of research, to improve the databases at our disposal, identifying and amending the critical issues. This work aims to provide the scientific community with a new sequence for the type strain *Paenisporosarcina quisquiliarum* SK 55, and to broaden the knowledge on the *Psychrobacillus psychrodurans* species, in particular considering the ancient strain Aquil_B6 found in an ancient Roman amphora.

7.2 Introduction

One of the most important and necessary practices in microbiology is the taxonomic identification of unknown bacterial strains. The 16S rRNA gene sequencing technique is one of the most used identification methods, that however shows several limitations. For example, for many phylogenetically close species, the potential of the 16S gene to provide taxonomic resolution at specie level is inadequate, being not able to uniquely and certainly identify the species of correspondence. This has been reported for the former genus *Bacillus* spp., where three reference strains shared greater than 99.5% similarity among the 16S rRNA gene sequence (1). A great similarity of this DNA region also characterizes the recently diverged species, thus making the distinction of several microbial species ineffective (2)(3). The quality of the sequences deposited in databases also plays a significant role in the results obtained. However, the reported large amount of low quality sequences deposited in the past years, as well as the estimated presence of errors or chimeras, may pose additional challenges in the identification procedures (4). Alternatively, the sequencing of other target genes, while capable of resolving some of these issues, is more time consuming. Nowadays Whole Genome Sequencing (WGS), thanks to its greater accessibility, is one of the most reliable and effective techniques for obtaining a unique and valid identification, as well as complete knowledge of the genetic characteristics of the investigated bacteria. WGS overcomes the issues associated with 16S rRNA gene sequencing due to the availability of methods based on the comparison of the entire DNA sequence that have been developed over time for more precise results. Since 1960, one of the most effective bacterial identification techniques has been DNA-DNA hybridization (DDH), efficient in providing stable and reproducible results. However due to the complexity of its execution, it has been gradually supplanted by new developed methods of comparison since the advent of the genomics era. These include the Average Nucleotide Identity (ANI), a metric based on the level of genomic similarity between the coding regions of two genomes (5). In addition to ANI, the increasingly affordability of genomic sequences have enabled the calculation of DDH *in silico* via the measurement of digital DNA – DNA hybridization (dDDH) (6). This made it possible to replace the complex laboratory operations necessary for the evaluation of DDH with simple and user friendly interface programs as for the free tool Genome-to-Genome Distance Calculator (GGDC) (7, 8). Based on this tool, the Type Strain Genome Server (TYGS) was developed, which can identify a query strain based on its entire genetic sequence by comparing dDDH values against an updated database of prokaryotic genomes of reference strains (9). However, even when using these tools, the quality of the databases used is critical. In fact, due to inconsistencies in the genetic sequences of some reference strains, using TYGS it was not possible to uniquely identify the belonging of a bacterial strain sequenced in a previous work (10). The unknown bacterial strain Aquil_B6, isolated together with 7

other bacilli from the content of an ancient Roman amphora of the 4th-5th century AD, in fact clustered for dDDH values both with the reference strain *Psychrobacillus psychrodurans* DSM 11713 and with the reference strain *Paenisporosarcina quisquiliarum* SK 55. Furthermore, TYGS showed that both reference strains could be identified as the same species. The sequences of these two strains, whose WGS Project is deposited on NCBI under the accession numbers GCA_900109875 and GCA_900114885, were reported as "anomalous assembly". To resolve this issue and obtain new correct genomes sequences to be deposited on NCBI database, for correctly and undoubtedly classifying strain Aquil_B6, the DNAs of both strains were re-sequenced. Due to the limited availability of other *Psychrobacillus psychrodurans* genomes, the sequencing of the *Psychrobacillus psychrodurans* DSM 30747 strain was also performed to provide the scientific community with new information on this specie. Nowadays, more and more species belonging to the genus *Psycrobacillus* are being studied for various peculiar characteristics, including the ability to degrade oils (11), produce bio-emulsifiers (12), and phosphate-solubilizing ability (13). These capabilities are also associated with the ubiquitous discovery of this genus, ranging from Egypt (14) to polar ices (12)(15), from ancient findings (10) to cleanrooms of space observatories (as reported for the DSM 30747 strain) or in clean-room environments of NASA (as reported in the bioproject PRJNA832800). The same considerations can be made for the genus *Paenisporosarcina*, an environmental bacterium of which many species were isolated in extreme environments (16). Given the small number of case studies on these species carried out especially in the last period, it is important to continue their studies, as their ability to resist adverse conditions could conceal important technological or industrial applications thanks to possible new metabolisms yet to be discovered (17)(18)(19)(20).

7.3 Results and Discussion

Analysis of the available genomes on NCBI database of the species *Psychrobacillus psychrodurans* and *Paenisporosarcina quisquiliarum*

During the preliminary identification of strain Aquil_B6 performed in a previous work (10), although a greater similarity with the reference strain *Psychrobacillus psychrodurans* DSM11713 (dDDH=88.8%) emerged, a match with dDDH values>70% also emerged with the reference strain *Paenisporosarcina quisquiliarum* SK 55. It was therefore decided to deepen this inconsistency in detail by verifying the correctness of the sequences deposited for these two reference strains. In the following figures and tables, the genomes of *Psychrobacillus psychrodurans* DSM11713 (assembly accession GCA_900114885, WGS project FOUN01) and *Paenisporosarcina quisquiliarum* SK 55 (assembly accession GCA_900109875, WGS project FOBQ01) already present in the TYGS database will be differentiated from the newly sequenced genomes by the indication *Old* (O) after their name. These genome sequences downloaded from NCBI, were used as query sequences to

obtain whole genome clustering (**Figure 1**), resulting in dDDH values reported in Table 1. As expected, the query sequences matched with the references strains, however FOUN01 and FOBQ01 genome sequences shared a dDDH value of 80% with a 0.04% difference in %GC. This dDDH value is above the threshold reported in literature (dDDH = 70) to consider two strains as separate species, suggesting that they belong to a single species (5)(21)(22). A further contradictory result was provided from the clustering based on the 16S rRNA gene sequence in Figure 2, where FOBQ01 sequence did not match with any deposited 16S rRNA gene sequence, without showing any match with the 16S rRNA gene sequence of *Paenisporosarcina quisquiliarum* SK 55 deposited with accession number DQ333897. These findings contradicted the taxonomic descriptions of these two distinct species (23)(24), and the absence of a match with the 16S rRNA gene sequence for *Paenisporosarcina quisquiliarum* SK 55 suggested that the FOBQ01 WGS project used as reference sequence by TYGS contained sequencing errors.

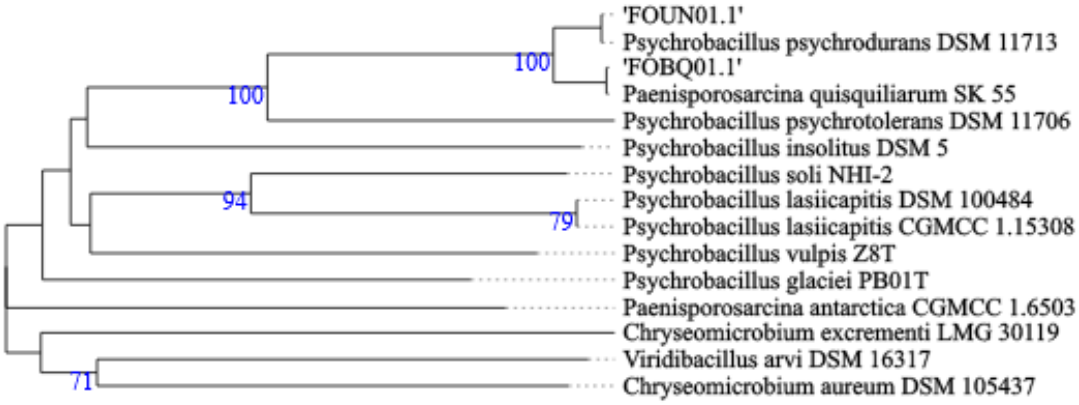


Figure 1, TYGS clusterization of FOUN01 and FOBQ01 WGS sequences

Table 1, dDDH calculation of FOUN01 and FOBQ01 sequences downloaded from NCBI trough TYGS

Query strain (NCBI)	Subject strain (TYGS)	dDDH (d_4, in %)	C.I. (d_4, in %)	G+C content difference (%)
FOUN01	*Psychrobacillus psychrodurans* DSM 11713 (O)	100.0	[100.0 - 100.0]	0.0
FOBQ01	*Paenisporosarcina quisquiliarum* SK 55 (O)	100.0	[100.0 - 100.0]	0.0
FOUN01	*Paenisporosarcina quisquiliarum* SK 55 (O)	80.0	[77.0 - 82.6]	0.04
FOBQ01	*Psychrobacillus psychrodurans* DSM 11713 (O)	80.0	[77.0 - 82.6]	0.04
FOBQ01	**FOUN01**	80.0	[77.0 - 82.6]	0.04

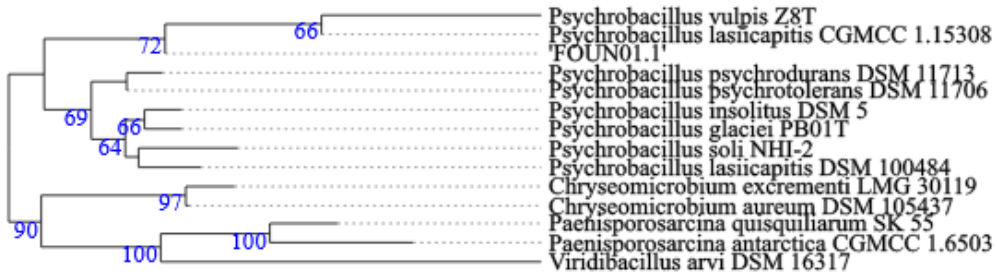


Figure 2, TYGS clusterization of FOUN01 and FOBQ01 16S rRNA gene sequence sequences

Genetic features of newly sequenced strains

The accession numbers for the Sequence Read Archive (SRA) and WGS of the newly assembled genomes analyzed in this study contained in bioproject PRJNA840842, as well as the previous *Psychrobacillus psychrodurans* Aquil_B6 from bioproject PRJNA811801, together with the total length and %GC comparison with previously available homologous strain sequences are reported in Table 2. It can be noted that *Psychrobacillus psychrodurans* DSM11713 and DSM30747 genomes were characterized by a similar length (4.03 and 4.06 Kbp, respectively) and %GC content (36.01-36.05%). When compared to the previously sequenced *Psychrobacillus psychrodurans* DSM11713 (O) genome, the new assembly showed a close length and %GC, suggesting the correctness of both sequencing. The genome of strain Aquil_B6 resulted very close to the reference strain, although it differed for a slightly longer length (4.26 Kbp) and a lower GC percentage (35.94%). The new genome assembly of *Paenisporosarcina quisquiliarum* SK 55, on the other hand, was much shorter (3.14 Kbp) and had a higher percentage of GC (39.71%) in respect to the deposited old sequence,

which appeared to be in contrast with the obtained results, showing a longer length of 4.03 Kbp (difference of 893465 bp) and a %GC difference of 3.75%. Therefore, the non-correspondence of the two strains under consideration is demonstrated by these values.

Table 2, general features, and accession numbers

Assembly	Total lenght	GC%	WGS accession	SRA accession
Newly assembled strains				
DSM11713	4027030	36,01	JAMKBK000000000	SRR19330377
DSM30747	4064800	36,05	JAMKBI000000000	SRR19330375
Aquil_B6	4256356	35,94	JAKXDZ000000000	SRR18190499
SK55	3140025	39,71	JAMKBJ000000000	SRR19330376
Previous reference strains				
SK55 (O)	4033490	35,96	FOBQ01000000	-
DSM11713 (O)	4016876	36,00	FOUN01000000	-

The assembly parameters and genetic characteristics of the strains under analysis were reported in Table 3. All genomes assembled with an L50 value between 5 and 6, with completeness values above 99.34% confirming the good results of the sequencing process. In addition to the differences in total length and %GC already analyzed above, the strains belonging to the *Psychrobacillus psychrodurans* species showed a greater number of genes, in the range of 3987 (for strain DSM 30747) and 4295 (for strain Aquil_B6), compared to *Paenisporosarcina quisquiliarum* SK 55, which presented 3203 genes.

Table 3, assembly statistics of the new assemblies

Assembly	Contigs	N50	L50	Completeness (%)	CDS	Gene	tRNA	tmRNA
DSM11713	57	232522	5	100	3945	4018	72	1
DSM30747	45	222761	6	99,34	3916	3987	70	1
Aquil_B6	79	253084	6	100	4224	4295	70	1
SK55	39	214672	5	99,34	3131	3203	71	1

Taxonomic analysis of new strains

Analyzing the clustering results of 16S rRNA gene sequence, predicted by TYGS, the three strains of *Psychrobacillus psychrodurans* DSM 11713, DSM30747, and Aquil_B6, although grouping together, did not show a perfect match with the 16S rRNA gene sequence of *Psychrobacillus psychrodurans* DSM 11713 already deposited with the accession number AJ277984, showing, on the contrary, a higher affinity with *Psychrobacillus psychrotolerans* DSM 11706 and *Psycrobacillus vulpis* Z8T (Figure 3). Once more this result suggests that the only comparison of 16S rRNA gene sequences between closely related species can lead to inaccurate identifications. Vice-versa, the predicted 16S rRNA gene sequence of *Paenisporosarcina quisquiliarum* SK 55 was correctly identified by clustering directly with the homologous reference deposited with the accession number DQ333897 (Figure 4).

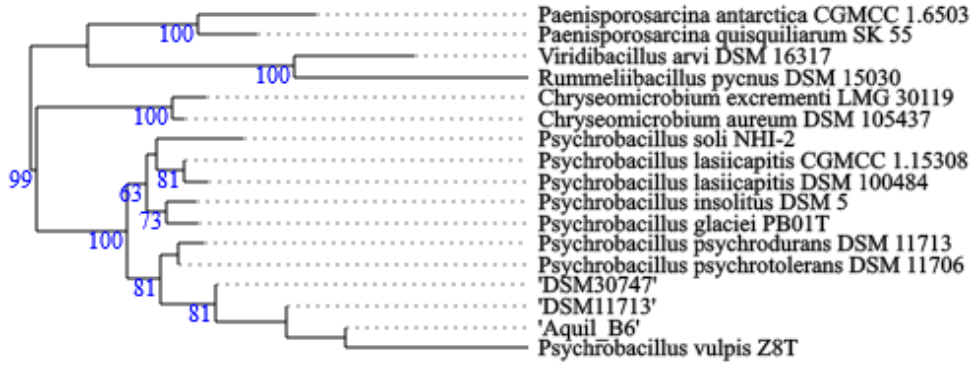


Figure 3, 16S rRNA gene sequence clustering of *Psychrobacillus psychrodurans* strains

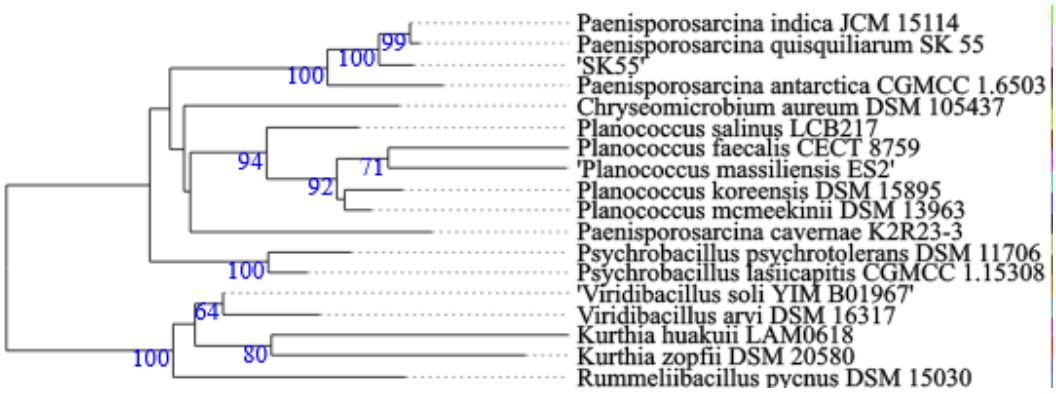


Figure 4, 16S rRNA gene sequence clustering of *Paenisporosarcina quisquiliarum* SK 55

Using the calculation of dDDH value on the whole genome sequence as comparison parameter, strains DSM 11713, DSM 30747, and Aquil_B6 clustered with the sequence FOUN01 of the reference strain *Psychrobacillus psychrodurans* DSM 11713, however together with the incorrect *Paenisporosarcina quisquiliarum* SK 55 sequence (**Figure 5**). The two assemblies of strain DSM 11713 had a dDDH

value of 100% based on the dDDH values reported in Table 4, confirming the accuracy of both sequences. Strains DSM 30747 and Aquil_B6 were also correctly attributed to this species, with dDDH values higher than the threshold limit, corresponding to 72% and 89%, respectively. It was therefore possible to identify with certainty Aquil_B6 strain as belonging to the species *Psychrobacillus psychrodurans.*

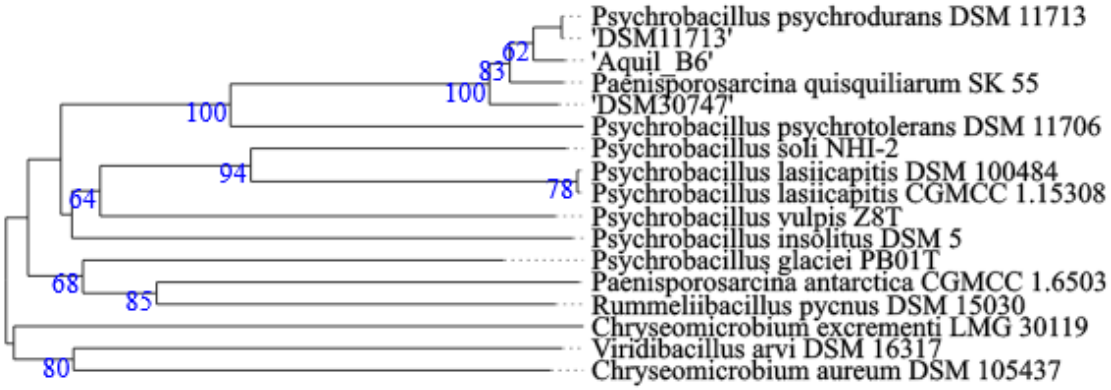


Figure 5, WGS clustering of *Psychrobacillus psychrodurans* strains

Table 4, dDDH values for *Psychrobacillus psychrodurans* strains

Query strain	Subject strain	dDDH (d_4, in %)	C.I. (d_4, in %)	G+C content difference (%)
DSM11713	*Psychrobacillus psychrodurans* DSM 11713 (O)	100.0	[100.0 - 100.0]	0.01
DSM30747	*Psychrobacillus psychrodurans* DSM 11713 (O)	71.8	[68.8 - 74.7]	0.05
Aquil_B6	*Psychrobacillus psychrodurans* DSM 11713 (O)	88.8	[86.4 - 90.9]	0.05
DSM11713	*Paenisporosarcina quisquiliarum* SK 55 (O)	80.0	[77.0 - 82.6]	0.05
DSM30747	*Paenisporosarcina quisquiliarum* SK 55 (O)	73.6	[70.5 - 76.4]	0.09
Aquil_B6	*Paenisporosarcina quisquiliarum* SK 55 (O)	79.0	[76.1 - 81.7]	0.01
DSM11713	**DSM30747**	71.9	[68.8 - 74.7]	0.04
Aquil_B6	**DSM30747**	71.9	[68.9 - 74.8]	0.11
Aquil_B6	**DSM11713**	88.8	[86.4 - 90.9]	0.07

On the contrary, the whole genome analysis of our new *Paenisporosarcina quisquiliarum* SK 55 sequence, resulted as a "new species" as observable in Table 5, with no matches above the dDDH threshold limit. The most similar species *Rummeliibacillus pycnus* DSM 15030 had dDDH similarity values of 24%, while with the previous sequence deposited for SK55 had a dDDH similarity of 20% with a %GC difference of 3.75. The same considerations can be derived from the observation of the tree proposed in Figure 6 where our sequence of the *Paenisporosarcina quisquiliarum* SK 55 strain did not seem to have some correspondence with any of the proposed reference strains. The same conclusions can be drawn from the Taxonomy Check made by NCBI on the assemblage (ASM2756331v1), where the closest species resulted *Paenisporosarcina indica* GCA_001939075.1 with an ANI similarity equal to 81.88%.

Table 5, dDDH values for *Paenisporosarcina quisquiliarum* SK 55

Query strain	Subject strain	dDDH (d_4, in %)	C.I. (d_4, in %)	G+C content difference (in %)
SK55'	*Paenisporosarcina quisquiliarum* SK 55 (O)	20.8	[18.6 - 23.2]	3.75
'SK55'	*Paenisporosarcina indica* JCM 15114	21.8	[19.6 - 24.3]	1.23
'SK55'	*Paenisporosarcina cavernae* K2R23-3	22.5	[20.3 - 25.0]	0.08
'SK55'	*Paenisporosarcina antarctica* CGMCC 1.6503	22.8	[20.5 - 25.2]	2.75
'SK55'	*Rummeliibacillus pycnus* DSM 15030	24.0	[21.7 - 26.5]	5.06

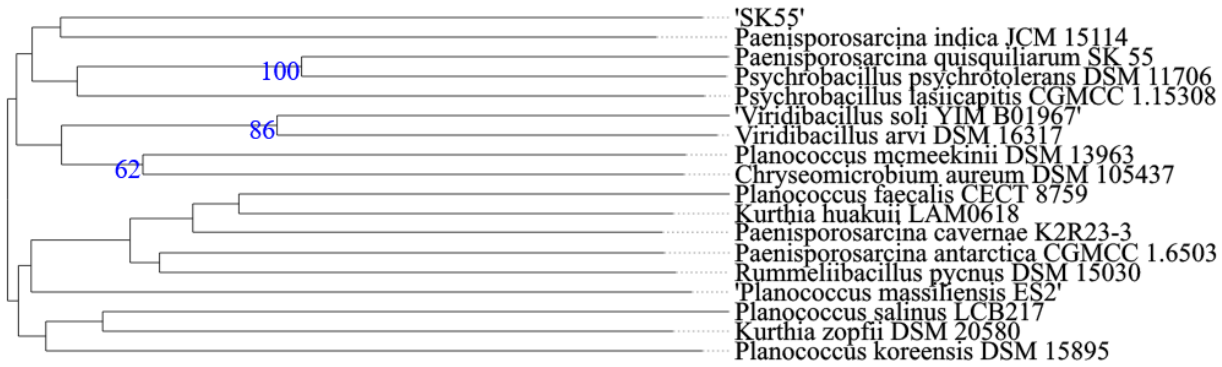


Figure 6, WGS clustering of *Paenisporosarcina quisquiliarum* SK 55

Strains characterization

The COG annotation in **Figure 7** shows how genes were distributed based on the attributed function, which was identified in 76.05% of cases. Most genes (36.62%) were linked to metabolic functions, in particular of amino acid transport and metabolism (E, 9.27%) followed by ion (P, 6.17%), carbohydrate (G, 5.78%), energy (C, 4.38%), lipid (I, 3.35%), coenzyme (H, 3.12%), nucleotide (F, 2.98%), and secondary metabolites (Q, 1.57 %) metabolisms. The remaining 20.15% were allocated to cellular processes and signaling, the most important of which were signal transduction mechanisms (T, 5.12%) and cell biogenesis (M, 4.57%). Finally, 19.28% of the functions were assigned to information storage and processing, mainly in transcription (K, 8.41%), translation (J, 5.79%) and replication recombination and repair (L, 5.05%). It can be observed that between the different strains a constant relationship of the different functions is maintained without significant variations.

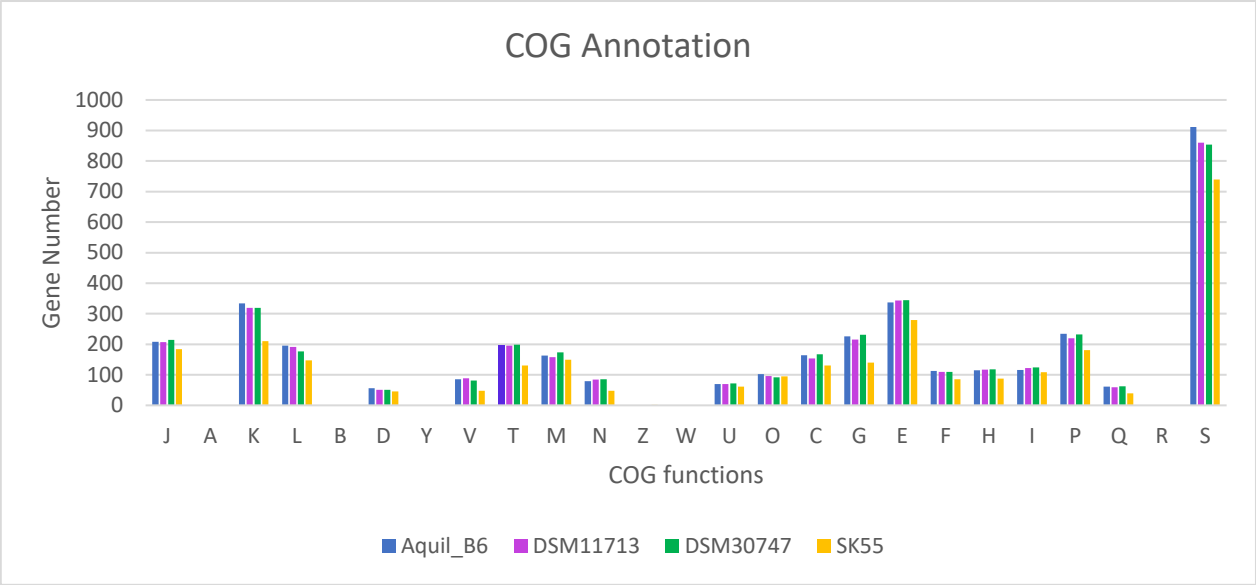


Figure 7, COG gene distribution

Using KEGG mapper, it was possible to reconstruct the metabolic pathways of the individual strains (**Supplementary Table 1**). All three analyzed *Psychrobacillus psychrodurans* strains shared the same complete pathways, with two exceptions: i) the first regards the metabolism of carbohydrates, in particular, strains Aquil_B6 and DSM 30747 shower the ability to synthesize UDP-galactose starting from galactose that wasn't predicted in the genome of the DSM 11713 strain; ii) regarding the polyketides biosynthesis, only strain DSM 30747 had a fully reconstructed dTDP-L-rhamnose biosynthesis pathway. *Paenisporosarcina quisquiliarum* SK55 distinguished from *Psychrobacillus psychrodurans* strains for the absence of several metabolisms. In carbohydrate metabolism, it lacked the Leloir pathway for the degradation of galactose, the biosynthesis of glycogen from glucose-1P, and for UDP-galactose from galactose. Also, in energy metabolism no F-type ATPase was identified, and in the nucleotide the metabolism for pyrimidine deoxyribonucleotide biosynthesis was missing. Moreover, the metabolism of amino acids presented incomplete pathways for the biosynthesis of threonine, methionine, valine/isoleucine and ornithine, and in the metabolism of cofactors and vitamins it was devoid of the thiamine salvage pathway, the biosynthesis of molybdenum cofactor and a pathway for pyridoxal P biosynthesis and of pantenoate. On the other hand, a complete pathway for formaldehyde assimilation was observed. Deoxyribonucleotide, lysine, and NAD biosynthesis, as well as C1 unit interconversion and C5 isoprenoid biosynthesis capability were found. Of particular attention, it was underlined the possible presence of a VraFG transporter resistance factor associated to antimicrobial peptides. Other features of the genomes were analyzed using specific tools described in the materials and methods section. Through functional annotation performed with Prokka, the possible presence of various bacterial cold-shock proteins, which confer resistance to low temperatures, was also identified (25). In all strains under examination the presence of CspA CspB

CspC and CspLA was in fact predicted. The presence of possible resistance factors or bacteriocins was not detected using RGI and BAGEL4. On the contrary, PHASTER analysis revealed the presence of a prophage in the genome of the strain DSM11713, which was fully identified as Paenibacillus phage PG1 (NC_021558). The analysis performed trough CRISPRCasFinder, however, revealed no possible related CRISPR-Cas defense systems. No strict matches were found in the research of possible resistance factors to biocides and heavy metals conducted using BacMet database. Only two possible resistance factors with an identity percentage greater than 80% were identified for strain Aquil_B6 (WP_063593029 nitrite reductase and WP_063593260 heavy metal translocating P-type ATPase).

Relation between Psychrobacillus psychrodurans strains

ANI values also confirmed the results of the dDDH analysis among the *Psychrobacillus psychrodurans* strains examined (**Table 6**). Despite the major difference in genome length and %GC, the reference strain DSM 11713 and the Aquil_B6 strain shared a dDDH similarity of 88.8% and ANI of 98.62%, putting them closer together than the DSM 30747 strain, which had dDDH value of 71.9% and ANI of 96.69% when compared to the reference strain.

Table 6, ANI and dDDH matrix for Psychrobacillus psychrodurans strains

		dDDH		
		Aquil_B6	**DSM 11713**	**DSM 30747**
ANI	**Aquil_B6**	-	88,8	71,9
	DSM 11713	98,629	-	71,9
	DSM 30747	96,65161	96,68996	-

Similar results were obtained by estimating the evolutionary distances of the genomes, and also in this case the DSM 11713 and Aquil_B6 strains were evolutionarily more similar at a temporal level than the DSM 30747 strain, which diverges more markedly (**Figure 8**).

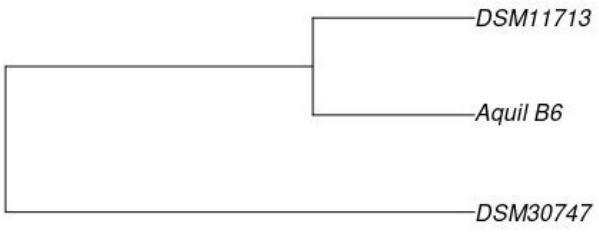


Figure 8, phylonium upgma evolutionary distance tree

A first comparative genome analysis of *Psychrobacillus psychrodurans* was shown in **Figure 9**. This assessment obtained by BRIG, in addition to the %GC representation, compared the genomes of strains Aquil_B6 and DSM 30747 to the reference strain DSM 11713 with a threshold value of 50%. It was possible to observe several points of differentiation between the reference and query genomes, also highlighting several differences between strains Aquil_B6 and DSM 30747.

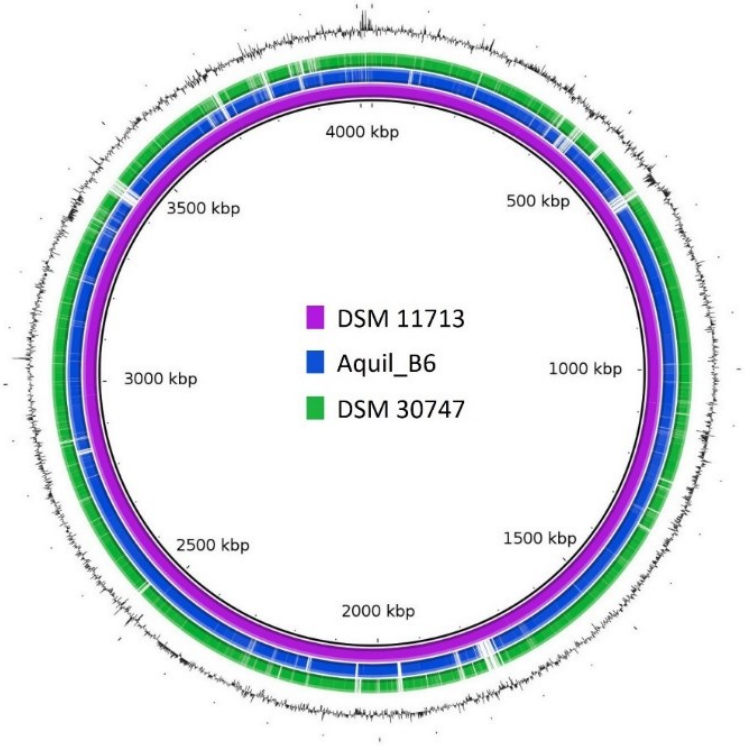


Figure 9, Genomes comparison obtained by BRIG of *Psychrobacillus psychrodurans* strains (circular graph).

The higher correlation between strains Aquil_B6 and DSM11713 was also highlighted by the analysis of the pangenome made with Roary, as reported in **Figure 10,** which depicted the distribution of genes in the different genomes. On a total of 5371 genes, a core genome of 3045 genes shared among all 3 genomes, and a shell genome shared from at least 2 genomes composed by 624 genes were identified. Strain Aquil_B6 was found to share a greater number of genes with strain DSM11713 (3467 genes) than with strain DSM 30747 (3137 genes). From the presence/absence table also obtained through Roary, 640 genes present exclusively in strain Aquil_B6, 673 in strain DSM 30747, and 389 in strain DSM 11713 were also predicted.

Figure 10, Roary gene matrix for *Psychrobacillus psychrodurans* strains

7.4 Conclusions

Through the resequencing of the genomes of *Psychrobacillus psychrodurans* DSM 11713 and *Paenisporosarcina quisquiliarum* SK 55 it was possible to identify troubles in the genomic sequence deposited for the latter. A new corrected reference genome for this species has therefore been provided. In the light of these results, it was possible to confirm the belonging of the Aquil_B6 strain to the *Psychrobacillus psychrodurans* species. The discovery of this microorganism in the internal content of a Roman amphora, which remained sealed for a long time confirms once again the resistance of this species to difficult environmental conditions as reported in the literature. The DSM 11713 strain was genetically the most similar to the Aquil_B6 strain, however, given the small number of genomes available, all sequenced in this work, it was not possible to make an effective comparison on the possible differences that such a prolonged isolation has produced on the strain. Anyway, fundamental steps have been taken on the knowledge of these species, which in the future can be expanded to understand the genetic basis of the resistance characteristics of these environmental organisms.

7.5 Materials and Methods

Bacterial strains and culture conditions

Strain Aquil_B6 was isolated as previously described from a 4th-5th century Roman amphora discovered still sealed in Aquileia (UD, Italy) (10). It was stored at -80° C at the University of Udine in Brain Heart Infusion (BHI) broth (Oxoid, Germany) added with 20% glycerol (Sigma, Germany). The other strains used in this study were obtained directly from the corresponding collections of microorganisms in freeze dried form: *Psychrobacillus psychrodurans* DSM 11713 and DSM 30747 *s*trains were obtained from Leibniz Institute DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen Collection (Germany), while *Paenisporosarcina quisquiliarum* strain SK 55 (JCM 14041) was ordered from Riken BRC, Microbe Division (JCM) (Japan). Strains were revitalized in BHI broth and their purity was verified on BHI agar (Oxoid, UK) streaked plates. Digital DNA-DNA hybridization (dDDH) and Average nucleotide identity (ANI). The reference draft genomes of *Paenisporosarcina quisquiliarum* strain SK 55 (GCA_900109875) and *Psychrobacillus psychrodurans* DSM 11713 (GCA_900114885) were used for genome comparison with the newly sequenced strains, as well as *Psychrobacillus psychrodurans* Aquil_B6 (GCA_022603175) from the NCBI database. Digital DNA hybridization (dDDH) calculation was performed using Type Strain Genome Server (TYGS) tool provided by Leibniz DSMZ Institute (9), while ANI values were calculated using FastANI (26) both used with default settings. A further comparison of the strains was done via JSpeciesWS database, using blastn (ANIb) and the Tetra Correlation Search (TCS) function (23).

DNA extraction and genome sequencing.

The DNA for genome sequencing was extracted from fresh cell culture growth overnight at 30 °C in BHI broth. Cells were pelleted by centrifugation at 5000 x g for 10 minutes. The DNA was extracted using the MagAttract HMW DNA kit (Qiagen, Germany) following the manufacturer's instructions. For genome sequencing, DNA library preparation was performed using the Nextera XT DNA sample preparation kit (Illumina, San Diego, CA, United States) according to the manufacturer's instructions. First, 1 ng input DNA from each sample was used for the library preparation which underwent fragmentation by sonication (BioRuptor-Diagenode, Belgium), adapter ligation, and amplification (Celero DNA-Seq kit, Tecan, Swiss). DNA sequencing was performed on a MiSeq instrument (Illumina) using a paired end 250 bp output sequencing Kit.

Genome assembly

The raw reads obtained from the sequencing process were carefully processed with the WGA-LP pipeline (27) using the following tools in default mode. Illumina adapters and quality trims were made with Trimmomatic v0.39 (28). FastQC v0.11.9 (29) was used to assess the quality of trimmed reads, and Kraken2 v2.0.8-b (30) was used to assess the possible presence of contaminants. Assembly was carried out using SPAdes v3.15.2 (31). The quality and completeness of the final assemblies was evaluated using CheckM v1.1.3 (32), Quast v5.0.2 (33), and SamTools v1.10 (34).

Genome annotation and characterization

The genomes were functionally annotated using Prokka 1.14.6 (35), reconstructing metabolisms and assigning COG annotation to identified proteins using EggNog (36) and classifying them according to the KEGG mapper function (37). Roary (38) was used in combination with BRIG to generate a BLASTN-based ring map for the analysis of gene distribution across genomes and the computation of the pangenome (39). PHASTER (40), CRISPRCasFinder (41), BAGEL4 (42) and RGI from CARDS (43) were also used to look for the presence of prophages, CRISPR-Cas systems, bacteriocins, and resistance factors. The BacMet database (44) was employed to identify antibacterial biocide and metal resistance genes. The estimation of evolutionary distances was made through phylonium (45).

7.6 References

1. Fox GE, Wisotzkey JD, Jurtshuk P. 1992. How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. Int J Syst Bacteriol 42:166–170.

2. Tang YW, Ellis NM, Hopkins MK, Smith DH, Dodge DE, Persing DH. 1998. Comparison of phenotypic and genotypic techniques for identification of unusual aerobic pathogenic gram-negative bacilli. J Clin Microbiol 36:3674–3679.

3. Enright MC, Carter PE, MacLean IA, McKenzie H. 1994. Phylogenetic relationships between some members of the genera Neisseria, Acinetobacter, Moraxella, and Kingella based on partial 16S ribosomal DNA sequence analysis. Int J Syst Bacteriol 44:387–391.

4. Janda JM, Abbott SL. 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls. J Clin Microbiol 45:2761–2764.

5. Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A 106:19126–19131.

6. Meier-Kolthoff JP, Klenk HP, Göker M. 2014. Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. Int J Syst Evol Microbiol 64:352–356.

7. Auch AF, Klenk HP, Göker M. 2010. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. Stand Genomic Sci 2:142–148.

8. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14.

9. Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun 10.

10. Colautti A, Comi G, De Paoli E, Peterlunger E, Novello M, Braidotti E, Pasini D, Iacumin L. 2022. Draft Genome Sequences of Eight Bacilli Isolated from an Ancient Roman Amphora. Microbiol Resour Announc 11:11–13.

11. Pham VHT, Jeong SW, Kim J. 2015. Psychrobacillus soli sp. nov., capable of degrading oil, isolated from oil-contaminated soil. Int J Syst Evol Microbiol 65:3046–3052.

12. da Silva MBF, da Mota FF, Jurelevicius D, de Carvalho Azevedo VA, da Costa MM, Góes-Neto A, Ramos RTJ, de Castro Soares S, Rosado AS, Seldin L. 2022. Genomic analyses of a novel bioemulsifier-producing Psychrobacillus strain isolated from soil of King George Island,

Antarctica. Polar Biol 45:691–701.

13. Chiba A, Peine M, Kublik S, Baum C. 2020. Complete Genome Sequence of Psychrobacillus sp . Strain Agricultural Soil in Germany 5:73–74.

14. Abd El-Rahman HA, Fritze D, Spröer C, Claus D. 2002. Two novel psychrotolerant species, Bacillus psychrotolerans sp. nov. and Bacillus psychrodurans sp. nov., which contain ornithine in their cell walls. Int J Syst Evol Microbiol 52:2127–2133.

15. Choi JY, Lee PC. 2020. Psychrobacillus glaciei sp. Nov., a psychrotolerant species isolated from an antarctic iceberg. Int J Syst Evol Microbiol 70:1947–1952.

16. Reddy GSN, Poorna Manasa B, Singh SK, Shivaji S. 2013. Paenisporosarcina indica sp. nov., a psychrophilic bacterium from a glacier, and reclassification of Sporosarcina antarctica Yu et al., 2008 as Paenisporosarcina antarctica comb. nov. and emended description of the genus Paenisporosarcina. Int J Syst Evol Microbiol 63:2927–2933.

17. Piotrowski Y, Gurung MK, Larsen AN. 2019. Characterization and engineering of a DNA polymerase reveals a single amino-acid substitution in the fingers subdomain to increase strand-displacement activity of A-family prokaryotic DNA polymerases. BMC Mol Cell Biol 20:1–11.

18. Kim Myung Kyum SE-H. WO 2021/006464 A1 - novel paenisporosarcina quisquiliarum 17mud 1-1541 strain and use thereof | The Lens.

19. Lee CW, Park SH, Lee SG, Park HH, Kim HJ, Park H, Park H, Lee JH. 2018. Crystal structure of dihydrodipicolinate reductase (PaDHDPR) from Paenisporosarcina sp. TG-14: Structural basis for NADPH preference as a cofactor. Sci Rep 8:1–12.

20. Doyle SM, Montross SN, Skidmore ML, Christner BC. 2013. Characterizing microbial diversity and the potential for metabolic function at -15 °c in the basal ice of Taylor Glacier, Antarctica. Biology (Basel) 2:1034–1053.

21. Auch AF, von Jan M, Klenk HP, Göker M. 2010. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2:117–134.

22. Ciufo S, Kannan S, Sharma S, Badretdin A, Clark K, Turner S, Brover S, Schoch CL, Kimchi A, DiCuccio M. 2018. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. Int J Syst Evol Microbiol 68:2386–2392.

23. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32:929–931.

24. Krishnamurthi S, Ruckmani A, Pukall R, Chakrabarti T. 2010. Psychrobacillus gen. nov. and proposal for reclassification of Bacillus insolitus Larkin & Stokes, 1967, B. psychrotolerans Abd-El Rahman et al., 2002 and B. psychrodurans Abd-El Rahman et al., 2002 as Psychrobacillus insolitus comb. nov., Psychrobacillus. Syst Appl Microbiol 33:367–373.

25. Ermolenko DN, Makhatadze GI. 2002. Bacterial cold-shock proteins. Cell Mol Life Sci 59:1902–1913.

26. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:1–8.

27. Rossi N, Colautti A, Iacumin L, Piazza C. 2021. WGA-LP: a pipeline for whole genome assembly of contaminated reads. Bioinformatics 1–3.

28. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

29. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

30. Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. Genome Biol 20:1–13.

31. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477.

32. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055.

33. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.

34. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin

R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.

35. Seemann T. 2014. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.

36. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Huerta-Cepas J. 2021. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. bioRxiv 5:2021.06.03.446934.

37. Kanehisa M, Sato Y. 2020. KEGG Mapper for inferring cellular functions from protein sequences. Protein Sci 29:28–35.

38. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693.

39. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. 2011. BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. BMC Genomics 12.

40. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–W21.

41. Couvin D, Bernheim A, Toffano-Nioche C, Touchon M, Michalik J, Néron B, Rocha EPC, Vergnaud G, Gautheret D, Pourcel C. 2018. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic Acids Res 46:W246–W251.

42. Van Heel AJ, De Jong A, Song C, Viel JH, Kok J, Kuipers OP. 2018. BAGEL4: A user-friendly web server to thoroughly mine RiPPs and bacteriocins. Nucleic Acids Res 46:W278–W281.

43. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen AL V., Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran HK, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar G V., McArthur AG. 2020. CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 48:D517–D525.

44. Pal C, Bengtsson-Palme J, Rensing C, Kristiansson E, Larsson DGJ. 2014. BacMet:

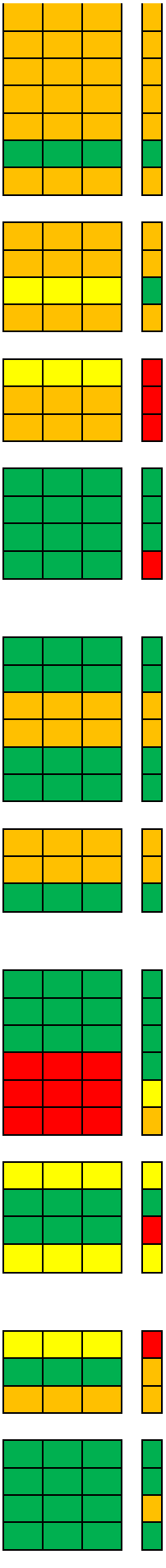
Antibacterial biocide and metal resistance genes database. Nucleic Acids Res 42:737–743.

45. Klötzl F, Haubold B. 2020. Phylonium: Fast estimation of evolutionary distances from large samples of similar genomes. Bioinformatics 36:2040–2046.

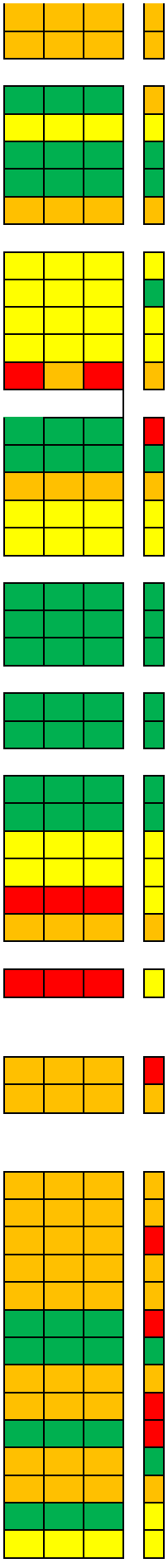
Supplementary materials

Supplementary Table 1, KEGG Mapper metabolisms

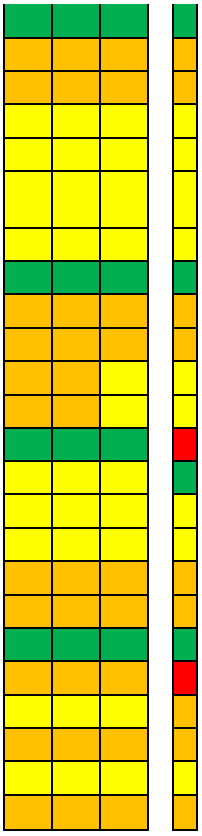
Strains				KEGG modules	Metabolism
Aquil_B6	DSM 30747	DSM 11713	SK55		
					Carbohydrate metabolism
					Central carbohydrate metabolism
				M00001	Glycolysis (Embden-Meyerhof pathway), glucose => pyruvate
				M00002	Glycolysis, core module involving three-carbon compounds
				M00003	Gluconeogenesis, oxaloacetate => fructose-6P
				M00307	Pyruvate oxidation, pyruvate => acetyl-CoA
				M00009	Citrate cycle (TCA cycle, Krebs cycle)
				M00010	Citrate cycle, first carbon oxidation, oxaloacetate => 2-oxoglutarate
				M00011	Citrate cycle, second carbon oxidation, 2-oxoglutarate => oxaloacetate
				M00004	Pentose phosphate pathway (Pentose phosphate cycle)
				M00006	Pentose phosphate pathway, oxidative phase, glucose 6P => ribulose 5P
				M00007	Pentose phosphate pathway, non-oxidative phase, fructose 6P => ribose 5P
				M00580	Pentose phosphate pathway, archaea, fructose 6P => ribose 5P
				M00005	PRPP biosynthesis, ribose 5P => PRPP
				M00008	Entner-Doudoroff pathway, glucose-6P => glyceraldehyde-3P + pyruvate
				M00308	Semi-phosphorylative Entner-Doudoroff pathway, gluconate => glycerate-3P
					Other carbohydrate metabolism
				M00014	Glucuronate pathway (uronate pathway)
				M00631	D-Galacturonate degradation (bacteria), D-galacturonate => pyruvate + D-glyceraldehyde 3P
				M00061	D-Glucuronate degradation, D-glucuronate => pyruvate + D-glyceraldehyde 3P
				M00632	Galactose degradation, Leloir pathway, galactose => alpha-D-glucose-1P
				M00552	D-galactonate degradation, De Ley-Doudoroff pathway, D-galactonate => glycerate-3P
				M00129	Ascorbate biosynthesis, animals, glucose-1P => ascorbate
				M00550	Ascorbate degradation, ascorbate => D-xylulose-5P
				M00854	Glycogen biosynthesis, glucose-1P => glycogen/starch
				M00855	Glycogen degradation, glycogen => glucose-6P
				M00565	Trehalose biosynthesis, D-glucose 1P => trehalose
				M00549	Nucleotide sugar biosynthesis, glucose => UDP-glucose
				M00554	Nucleotide sugar biosynthesis, galactose => UDP-galactose
				M00892	UDP-N-acetyl-D-glucosamine biosynthesis, eukaryotes, glucose => UDP-GlcNAc
				M00909	UDP-N-acetyl-D-glucosamine biosynthesis, prokaryotes, glucose => UDP-GlcNAc
				M00012	Glyoxylate cycle
				M00373	Ethylmalonyl pathway
				M00740	Methylaspartate cycle
				M00532	Photorespiration
				M00013	Malonate semialdehyde pathway, propanoyl-CoA => acetyl-CoA
				M00741	Propanoyl-CoA metabolism, propanoyl-CoA => succinyl-CoA
				M00131	Inositol phosphate metabolism, Ins(1,3,4,5)P4 => Ins(1,3,4)P3 => myo-inositol
					Energy metabolism
					Carbon fixation
				M00165	Reductive pentose phosphate cycle (Calvin cycle)
				M00166	Reductive pentose phosphate cycle, ribulose-5P => glyceraldehyde-3P
				M00167	Reductive pentose phosphate cycle, glyceraldehyde-3P => ribulose-5P
				M00168	CAM (Crassulacean acid metabolism), dark
				M00170	C4-dicarboxylic acid cycle, phosphoenolpyruvate carboxykinase type



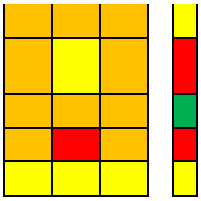
M00173	Reductive citrate cycle (Arnon-Buchanan cycle)
M00376	3-Hydroxypropionate bi-cycle
M00375	Hydroxypropionate-hydroxybutylate cycle
M00374	Dicarboxylate-hydroxybutyrate cycle
M00377	Reductive acetyl-CoA pathway (Wood-Ljungdahl pathway)
M00579	Phosphate acetyltransferase-acetate kinase pathway, acetyl-CoA => acetate
M00620	Incomplete reductive citrate cycle, acetyl-CoA => oxoglutarate
	Methane metabolism
M00357	Methanogenesis, acetate => methane
M00346	Formaldehyde assimilation, serine pathway
M00345	Formaldehyde assimilation, ribulose monophosphate pathway
M00344	Formaldehyde assimilation, xylulose monophosphate pathway
	Nitrogen metabolism
M00530	Dissimilatory nitrate reduction, nitrate => ammonia
M00529	Denitrification, nitrate => nitrogen
M00804	Complete nitrification, comammox, ammonia => nitrite => nitrate
	ATP synthesis
M00151	Cytochrome bc1 complex respiratory unit
M00155	Cytochrome c oxidase, prokaryotes
M00416	Cytochrome aa3-600 menaquinol oxidase
M00157	F-type ATPase, prokaryotes and chloroplasts
	Lipid metabolism
	Fatty acid metabolism
M00082	Fatty acid biosynthesis, initiation
M00083	Fatty acid biosynthesis, elongation
M00873	Fatty acid biosynthesis in mitochondria, animals
M00874	Fatty acid biosynthesis in mitochondria, fungi
M00086	beta-Oxidation, acyl-CoA synthesis
M00087	beta-Oxidation
	Lipid metabolism
M00088	Ketone body biosynthesis, acetyl-CoA => acetoacetate/3-hydroxybutyrate/acetone
M00089	Triacylglycerol biosynthesis
M00093	Phosphatidylethanolamine (PE) biosynthesis, PA => PS => PE
	Nucleotide metabolism
	Purine metabolism
M00048	Inosine monophosphate biosynthesis, PRPP + glutamine => IMP
M00049	Adenine ribonucleotide biosynthesis, IMP => ADP,ATP
M00050	Guanine ribonucleotide biosynthesis, IMP => GDP,GTP
M00053	Deoxyribonucleotide biosynthesis
M00958	Adenine ribonucleotide degradation, AMP => Urate
M00959	Guanine ribonucleotide degradation, GMP => Urate
	Pyrimidine metabolism
M00051	Uridine monophosphate biosynthesis, glutamine (+ PRPP) => UMP
M00052	Pyrimidine ribonucleotide biosynthesis, UMP => UDP/UTP,CDP/CTP
M00053	Pyrimidine deoxyribonuleotide biosynthesis, CDP => dCTP
M00938	Pyrimidine deoxyribonuleotide biosynthesis, UDP => dTTP
	Amino acid metabolism
	Serine and threonine metabolism
M00020	Serine biosynthesis, glycerate-3P => serine
M00018	Threonine biosynthesis, aspartate => homoserine => threonine
M00033	Ectoine biosynthesis, aspartate => ectoine
	Cysteine and methionine metabolism
M00021	Cysteine biosynthesis, serine => cysteine
M00609	Cysteine biosynthesis, methionine => cysteine
M00017	Methionine biosynthesis, apartate => homoserine => methionine
M00034	Methionine salvage pathway



M00035	Methionine degradation
M00368	Ethylene biosynthesis, methionine => ethylene
	Branched-chain amino acid metabolism
M00019	Valine/isoleucine biosynthesis, pyruvate => valine / 2-oxobutanoate => isoleucine
M00535	Isoleucine biosynthesis, pyruvate => 2-oxobutanoate
M00570	Isoleucine biosynthesis, threonine => 2-oxobutanoate => isoleucine
M00432	Leucine biosynthesis, 2-oxoisovalerate => 2-oxoisocaproate
M00036	Leucine degradation, leucine => acetoacetate + acetyl-CoA
	Lysine metabolism
M00016	Lysine biosynthesis, succinyl-DAP pathway, aspartate => lysine
M00525	Lysine biosynthesis, acetyl-DAP pathway, aspartate => lysine
M00526	Lysine biosynthesis, DAP dehydrogenase pathway, aspartate => lysine
M00527	Lysine biosynthesis, DAP aminotransferase pathway, aspartate => lysine
M00030	Lysine biosynthesis, AAA pathway, 2-oxoglutarate => 2-aminoadipate => lysine
	Arginine and proline metabolism
M00028	Ornithine biosynthesis, glutamate => ornithine
M00844	Arginine biosynthesis, ornithine => arginine
M00845	Arginine biosynthesis, glutamate => acetylcitrulline => arginine
M00029	Urea cycle
M00015	Proline biosynthesis, glutamate => proline
	Polyamine biosynthesis
M00133	Polyamine biosynthesis, arginine => agmatine => putrescine => spermidine
M00134	Polyamine biosynthesis, arginine => ornithine => putrescine
M00135	GABA biosynthesis, eukaryotes, putrescine => GABA
	Histidine metabolism
M00026	Histidine biosynthesis, PRPP => histidine
M00045	Histidine degradation, histidine => N-formiminoglutamate => glutamate
	Aromatic amino acid metabolism
M00022	Shikimate pathway, phosphoenolpyruvate + erythrose-4P => chorismate
M00023	Tryptophan biosynthesis, chorismate => tryptophan
M00024	Phenylalanine biosynthesis, chorismate => phenylpyruvate => phenylalanine
M00025	Tyrosine biosynthesis, chorismate => HPP => tyrosine
M00533	Homoprotocatechuate degradation, homoprotocatechuate => 2-oxohept-3-enedioate
M00038	Tryptophan metabolism, tryptophan => kynurenine => 2-aminomuconate
	Other amino acid metabolism
M00027	GABA (gamma-Aminobutyrate) shunt
	Glycan metabolism
	Lipopolysaccharide metabolism
M00064	ADP-L-glycero-D-manno-heptose biosynthesis
M00922	CMP-Neu5Ac biosynthesis
	Metabolism of cofactors and vitamins
	Cofactor and vitamin metabolism
M00127	Thiamine biosynthesis, prokaryotes, AIR (+ DXP/tyrosine) => TMP/TPP
M00895	Thiamine biosynthesis, prokaryotes, AIR (+ DXP/glycine) => TMP/TPP
M00896	Thiamine biosynthesis, archaea, AIR (+ NAD+) => TMP/TPP
M00897	Thiamine biosynthesis, plants, AIR (+ NAD+) => TMP/thiamine/TPP
M00898	Thiamine biosynthesis, pyridoxal-5P => TMP/thiamine/TPP
M00899	Thiamine salvage pathway, HMP/HET => TMP
M00125	Riboflavin biosynthesis, plants and bacteria, GTP => riboflavin/FMN/FAD
M00911	Riboflavin biosynthesis, fungi, GTP => riboflavin/FMN/FAD
M00124	Pyridoxal-P biosynthesis, erythrose-4P => pyridoxal-P
M00916	Pyridoxal-P biosynthesis, R5P + glyceraldehyde-3P + glutamine => pyridoxal-P
M00115	NAD biosynthesis, aspartate => quinolinate => NAD
M00912	NAD biosynthesis, tryptophan => quinolinate => NAD
M00119	Pantothenate biosynthesis, valine/L-aspartate => pantothenate
M00913	Pantothenate biosynthesis, 2-oxoisovalerate/spermine => pantothenate



M00120	Coenzyme A biosynthesis, pantothenate => CoA
M00914	Coenzyme A biosynthesis, archaea, 2-oxoisovalerate => 4-phosphopantoate => CoA
M00572	Pimeloyl-ACP biosynthesis, BioC-BioH pathway, malonyl-ACP => pimeloyl-ACP
M00881	Lipoic acid biosynthesis, plants and bacteria, octanoyl-ACP => dihydrolipoyl-E2/H
M00882	Lipoic acid biosynthesis, eukaryotes, octanoyl-ACP => dihydrolipoyl-H
M00883	Lipoic acid biosynthesis, animals and bacteria, octanoyl-ACP => dihydrolipoyl-H => dihydrolipoyl-E2
M00884	Lipoic acid biosynthesis, octanoyl-CoA => dihydrolipoyl-E2
M00126	Tetrahydrofolate biosynthesis, GTP => THF
M00840	Tetrahydrofolate biosynthesis, mediated by ribA and trpF, GTP => THF
M00841	Tetrahydrofolate biosynthesis, mediated by PTPS, GTP => THF
M00842	Tetrahydrobiopterin biosynthesis, GTP => BH4
M00843	L-threo-Tetrahydrobiopterin biosynthesis, GTP => L-threo-BH4
M00880	Molybdenum cofactor biosynthesis, GTP => molybdenum cofactor
M00140	C1-unit interconversion, prokaryotes
M00141	C1-unit interconversion, eukaryotes
M00846	Siroheme biosynthesis, glutamyl-tRNA => siroheme
M00868	Heme biosynthesis, animals and fungi, glycine => heme
M00121	Heme biosynthesis, plants and bacteria, glutamate => heme
M00926	Heme biosynthesis, bacteria, glutamyl-tRNA => coproporphyrin III => heme
M00847	Heme biosynthesis, archaea, siroheme => heme
M00122	Cobalamin biosynthesis, cobyrinate a,c-diamide => cobalamin
M00117	Ubiquinone biosynthesis, prokaryotes, chorismate (+ polyprenyl-PP) => ubiquinol
M00116	Menaquinone biosynthesis, chorismate (+ polyprenyl-PP) => menaquinol
M00932	Phylloquinone biosynthesis, chorismate (+ phytyl-PP) => phylloquinol
	Biosynthesis of terpenoids and polyketides
	Terpenoid backbone biosynthesis
M00095	C5 isoprenoid biosynthesis, mevalonate pathway
M00849	C5 isoprenoid biosynthesis, mevalonate pathway, archaea
M00096	C5 isoprenoid biosynthesis, non-mevalonate pathway
M00364	C10-C20 isoprenoid biosynthesis, bacteria
M00365	C10-C20 isoprenoid biosynthesis, archaea
M00366	C10-C20 isoprenoid biosynthesis, plants
M00367	C10-C20 isoprenoid biosynthesis, non-plant eukaryotes
	Polyketide sugar unit biosynthesis
M00793	dTDP-L-rhamnose biosynthesis
	Biosynthesis of other secondary metabolites
	Biosynthesis of phytochemical compounds
M00039	Monolignol biosynthesis, phenylalanine/tyrosine => monolignol
M00942	Pterocarpan biosynthesis, daidzein => medicarpin
	Biosynthesis of other antibiotics
M00877	Kanosamine biosynthesis glucose 6-phosphate => kanosamine
M00787	Bacilysin biosynthesis, prephenate => bacilysin
	Xenobiotics biodegradation
	Aromatics degradation
M00568	Catechol ortho-cleavage, catechol => 3-oxoadipate
M00569	Catechol meta-cleavage, catechol => acetyl-CoA / 4-methylcatechol => propanoyl-CoA
M00878	Phenylacetate degradation, phenylaxetate => acetyl-CoA/succinyl-CoA
	Signature modules
	Gene set
	Pathogenicity
M00860	Bacillus anthracis pathogenicity signature, polyglutamic acid capsule biosynthesis
	Drug resistance
M00625	Methicillin resistance
M00627	beta-Lactam resistance, Bla system
M00704	Tetracycline resistance, efflux pump Tet38



M00725	Cationic antimicrobial peptide (CAMP) resistance, dltABCD operon
M00726	Cationic antimicrobial peptide (CAMP) resistance, lysyl-phosphatidylglycerol (L-PG) synthase MprF
M00730	Cationic antimicrobial peptide (CAMP) resistance, VraFG transporter
M00769	Multidrug resistance, efflux pump MexPQ-OpmE
M00700	Multidrug resistance, efflux pump AbcA

Module set	
Metabolic capacity	
M00618	Acetogen
M00615	Nitrate assimilation

Legend:

- Complete
- Incomplete for 1 step
- Incomplete for >1 step
- Absent

Chapter 8: In search of ancient bacteria: bacilli isolated from an ancient Roman amphora of the IV-V century AD

Andrea Colautti, Giuseppe Comi, Enrico Peterlunger, Lucilla Iacumin

Department of Agricultural, Food, Environmental and Animal Science, University of Udine, Udine, Italy

8.1 Abstract

In this study, the microbiological content of an ancient Roman amphora dating back to the 4th-5th centuries AD was analyzed. A high load of still viable microbial forms identified as sporogenic bacilli was discovered inside the intact and sealed amphora. These bacteria, thanks to the resistance of their spores, can withstand harsh conditions and survive for long periods of time. Eight isolates were selected for further genetic characterization after DNA extraction. Thanks to the new technologies of Whole Genome Sequencing, it is in fact possible to perform a very precise taxonomic identification as well as a more in-depth and extensive genetic characterization. Using ANI, dDDH, and %GC values as classification methods, 5 isolates were classifiable as *L. fusiformis*, *B. muralis*, *B. psychrodurans*, and *B. frigoritolerans*, but 3 isolates were not certainly identifiable as known species. These differences could indicate an evolutionary separation from the current bacilli or the identification of new microbial species. The possible metabolisms present in the eight strains, as well as the presence of insertion sequences, prophages, CRISPR-Cas systems, and bacteriocins, were also described using bioinformatics tools. To compare this genetic trait with current bacteria, special attention was paid to the identification of resistance and virulence factors, which revealed the possible presence of several resistance genes.

Importance

The study of ancient bacteria and their genomic features, which have remained unchanged for centuries, can shed light on several bacterial evolutionary phenomena. For example, by comparing the spread of antibiotic resistance factors, it is possible to determine whether this characteristic is a result of today's widespread use of antibiotics or an inherent feature of bacteria. Furthermore, the isolation of these bacteria allows us to study the genetic characteristics of Roman-era food-contaminating bacteria.

Keywords: Paleomicrobiology, genome-sequencing, antibiotic-resistance

8.2 Introduction

The discovery of the amphora analyzed in this work took place in Aquileia (Udine, Italy) (45°45'05.9"N 13°21'03.8"E) during archaeological excavations made in the second-third decade of 1900. It was kept in custody at the amphorae warehouse of the Archaeological Museum of Aquileia until the moment of opening. The study of these amphorae allows us to understand and also reconstruct the economic importance of the area in the ancient Roman period, together with the uses and habits of the population at the time. The various sources show us that from the 1st to the 4th century AD, Aquileia was a rich and important city, at the center of a continuous and wide commercial flow of different origins, as evidenced by the amphorae themselves and their contents. From the study of the amphorae, their typology, and their content, it is possible to describe an Aquileia that was at the center of a lively trade with the role of both importer and exporter from and to the main cities of the Adriatic, Greece, Africa and from Spain. The setting of wine was relevant during the first and middle imperial ages from Gallia and from Tarraconese. These were regions that exported large quantities of wine, the quality of which was lower than wine from other regions. Subsequently, the import seems to be limited to Greece, known for producing quality wine for the more affluent classes. This leads to formulate the hypothesis that starting from the second century, the production of local wine has been implemented, achieving the satisfaction of internal needs (14). From a microbiological point of view, studies on ancient genetic material can help in understanding how microorganisms and their metabolisms evolved over time. Starting with Seaward (1976) (1), who discovered *Thermoactinomyces vulgaris* spores in organic material dating back to 85-125 BC, several works have been based on the analysis of ancient bacterial genomes addressing various aspects, such as the virulence of ancient pathogens (2)(3), the ancient human microbiome (4)(5)(6)(7), and the study of anthropological dynamics (8). *Bacillus* spp. is one of the species that has been reported to be capable to survive in a dormant state for an extended period of time. This feature is allowed by the ability of this microbial species to produce spores that can withstand a variety of extreme factors including radiations, chemicals, and heat (9). For example, viable forms of *Bacillus* spp. were discovered during the examination of a mummy still sealed in a sarcophagus (10). Furthermore, vital spores belonging to various microorganisms, including bacteria identified as *B. subtilis*, *B. cereus*, and *B. megaterium*, were discovered in soil samples recovered from archaeological excavations dated 1053 BC and 700-900 BC (11). These ancient bacteria and their genetic material can provide valuable information. By evaluating the presence of resistance factors in these genomes, a worthful clue on the evolution of the current antibiotic-resistance issue can be discovered. This phenomenon is thought to be becoming more common as antibiotics have become more widely used in recent years. However, the presence of numerous resistance factors identified in genomic material recovered from ancient sediments

contradicts this hypothesis (12). Similar evidence was discovered in the genomes of 28 bacteria that had been preserved in Siberian permafrost for 2.7 million years. When compared to current bacteria, their antibiotic resistance profile was identical (13). In this study, 8 bacilli strains isolated from the content of an ancient Roman amphora, found still intact and sealed, dating back to the IV-V century AD were studied.

The aim of the study was the taxonomic identification and the in-depth genetic study of the isolated bacilli. To obtain a precise classification of these still viable microorganisms, their genomes were annotated and characterized to understand more about their metabolisms and genetic characteristics. Particular attention was paid to the presence of virulence genes, in order to confirm the presence of these genes in bacilli that lived during the time of ancient Rome and compare their spread to that of modern homologous sequenced bacteria. Furthermore, because several studies have shown that these microorganisms have high resistance to metals and toxic substances (15)(16)(17)(18)(19) resistance genes for these factors were searched. This research provided us with an overview of the characteristics of these ancient bacteria, with the goal of determining whether they belonged to unidentified species and what metabolic differences they had with modern era bacteria.

8.3 Results

Characteristics of the amphora

The amphora under study was a Byzacena amphora of African origin (**Figure 1**). These amphorae constitute a group of multiple typological variants, generally known as "cylindrical of medium size" and are widely documented in Aquileia. It is assumed that they generally carried garum (fish sauce) and, perhaps, also oil, throughout the 4th century and the first decades of the 5th century.

Figure 1, image of the amphora from the warehouse of the Archaeological Museum of Aquileia

In the specific case, the opening of the amphora brought to light a very heterogeneous dehydrated organic material that excluded the content may have been oil **(Figure 2**). The preliminary analysis of

the organic material determined the presence of different plant leaves (thyme, rosemary, juniper), grape seeds and grapes. The absence of fish remains, NaCl and glutamate suggested that the content was not even *garum*, proposing the hypothesis that it contained wine. However, the study on the remains found is ongoing.

Figure 2. Material retrieved from the amphora. The first plate contains the bulk material constituted by dried leaves and seeds observable in more detail in the second plate. In the third plate it is observable an entire dried grape retrieved from the inside.

Microbiological Sampling and sporulation test

From the sampling of the residues contained in the amphora, based on the morphological characteristics of the colonies grown on the different mediums, it was possible to observe only colonies attributable to *Bacillus* spp., while the growth of other species was not evidenced by the different selective media used. *Bacillus* spp. colonies were counted on BHI medium, and resulted present at a concentration of 7.26±0.09 log CFU/g. A total of 25 colonies were isolated independently from their morphology and size, and profiled by RAPD, Rep-PCR, SAU-PCR to eliminate clones. Among the 25 strains, 8 different genetic profiles were defined, and represented through the dendrograms shown in **Figure 3**.

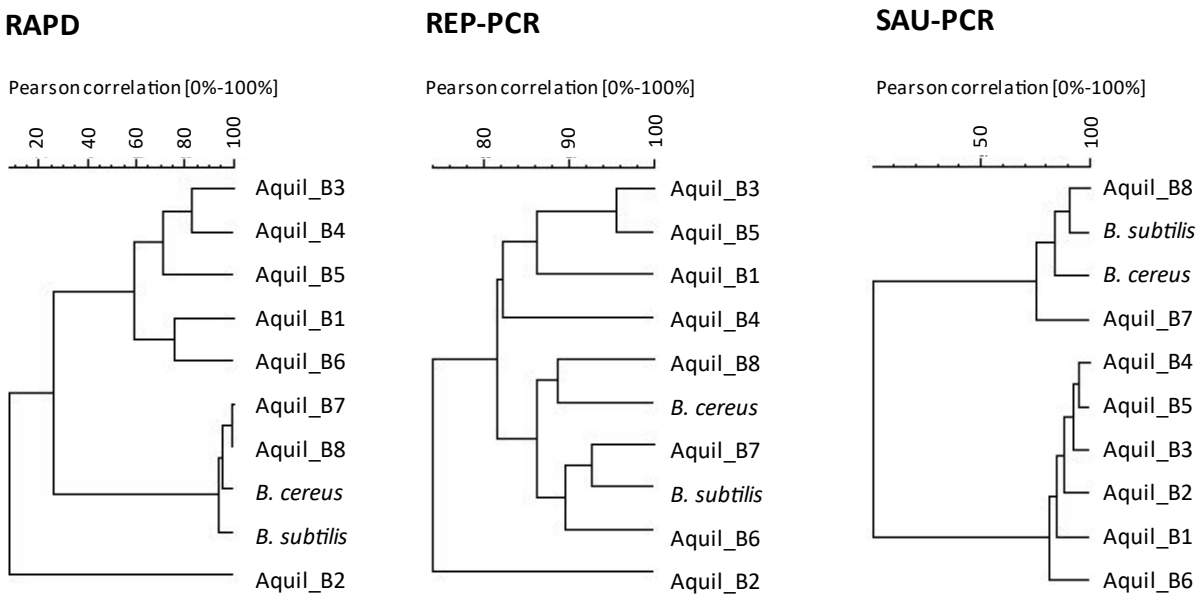


Figure 3, Cluster analysis of the profiles obtained from the 8 different *Bacillus* spp. strains isolated from the amphora. Calculation of similarity in the profiles of bands was based on Pearson product-moment correlation coefficients. Dendrograms were obtained by means of the Unweighted Pair Group Method using Arithmetic Average (UPGMA) clustering algorithms.

The sequencing of about 600 bp of the 16S rRNA gene allowed a first identification of the eight different strains further subjected to whole genome sequencing. **Table 1** reports the accession number (AN) for the 16S rRNA partial gene sequence, SRA and GeneBank of the genomes.

Table 1, Deposited accession numbers

Assembly	GeneBank AN	SRA AN	16S RNA AN
Aquil_B1	JAKXEE000000000	SRR18190504	ON326590
Aquil_B2	JAKXED000000000	SRR18190503	ON326591
Aquil_B3	JAKXEC000000000	SRR18190502	ON326592
Aquil_B4	JAKXEB000000000	SRR18190501	ON326593
Aquil_B5	JAKXEA000000000	SRR18190500	ON326594
Aquil_B6	JAKXDZ000000000	SRR18190499	ON326595
Aquil_B7	JAKXDY000000000	SRR18190498	ON326596
Aquil_B8	JAKXDX000000000	SRR18190497	ON326597

Following the staining with malachite green and observation under the microscope, the sporulation ability of all the bacteria considered was confirmed.

General features and identification

The general features of the obtained assemblies were summarized in **Table 2**. The 8 genomes assembled with an average value of L50 = 4 and N50 = 572788, with an average length of 5.32Mb,

ranging between 6.68 Mbp of strain B4 and 4.26 Mbp of strain B6. CheckM confirmed a good result of the sequencing, reporting a level of completeness above 98% for all strains. The %GC differed among the strains with values ranging between 35.94% of strain B6 and 41.27% of strain B3, with a mean value of 39.53%. On average, 5248 genes were identified by Prokka, ranging between 6727 genes of strain B4 and 4295 of strain B6.

Table 2, statistics of assembled genomes

	Aquil_B1	Aquil_B2	Aquil_B3	Aquil_B4	Aquil_B5	Aquil_B6	Aquil_B7	Aquil_B8
Total contigs	25	34	38	68	38	79	46	78
Total Length[a]	5649653	4643302	5057074	6677279	5067063	4256356	5521551	5654249
GC (%)[a]	40.24	37.49	41.26	39.51	41.27	35.94	40.32	40.24
N50[a]	910271	1011198	641321	317977	641321	253084	613520	193610
CDS[b]	5475	4578	4784	6633	4784	4224	5360	5474
Genes[b]	5558	4665	4869	6727	4869	4295	5445	5557
Completeness (%)[c]	98.91	99.93	98.77	98.91	98.91	100	98.91	98.91

[a] Determined using Quast
[b] Determined using Prokka
[c] Determined using CheckM

A preliminary analysis was performed using TYGS, comparing the 8 genomes with all the reference strains in this database using dDDH values to obtain a first taxonomic identification (**Figure 4**). Following this first clue, ANI and dDDH values (**Supplementary Table 1**) were calculated between the analyzed strains and the strains deposited on NCBI for the most genetically related species (**Table Supplementary 2**).

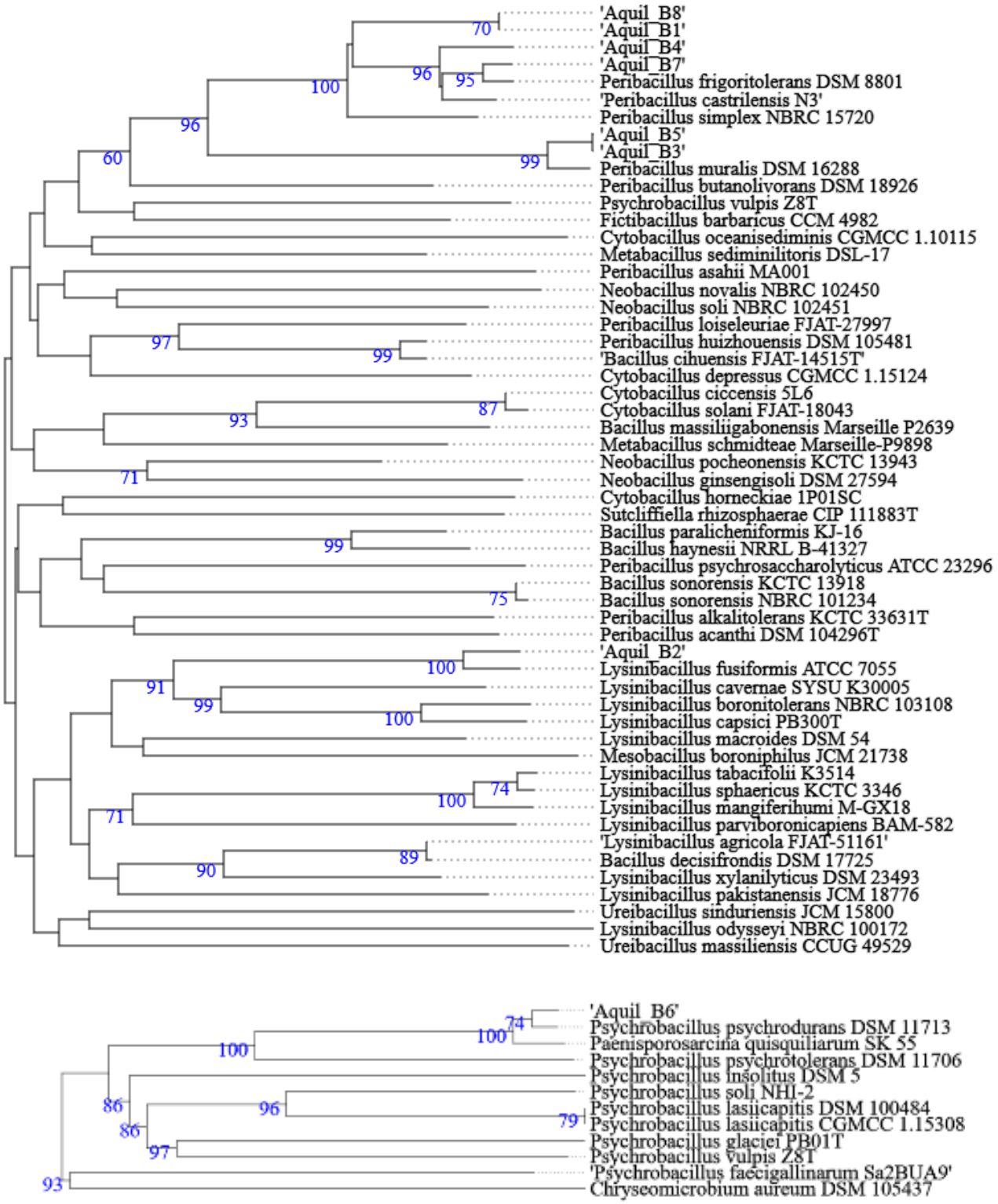


Figure 4, WGS clusterization with reference strains based on dDDH of the analyzed strains

Analyzing in detail, strain Aquil_B2 was characterized by a 4.64 Mbp genome with a %GC = 37.49 and 4665 predicted genes. It clustered with the type strain *Lysinbacillus fusiformis* ATCC7055 with dDDH values of 77.00, ANI of 97.10 and a GC difference of 0.16%, with values above the ANI and

dDDH threshold with most of the genomes deposited for this species on NCBI, thus identifying it with certainty as belonging to the species *Lysinbacillus fusiformis*. Analyzing the genes mapping (**Figure 5**), it can be observed a core-genome shared among all the sequences of this species, in which, strain Aquil_B2 clustered closer to 4 other strains (SW-B9, ATCC 55673, Cu1_5 and RB-21).

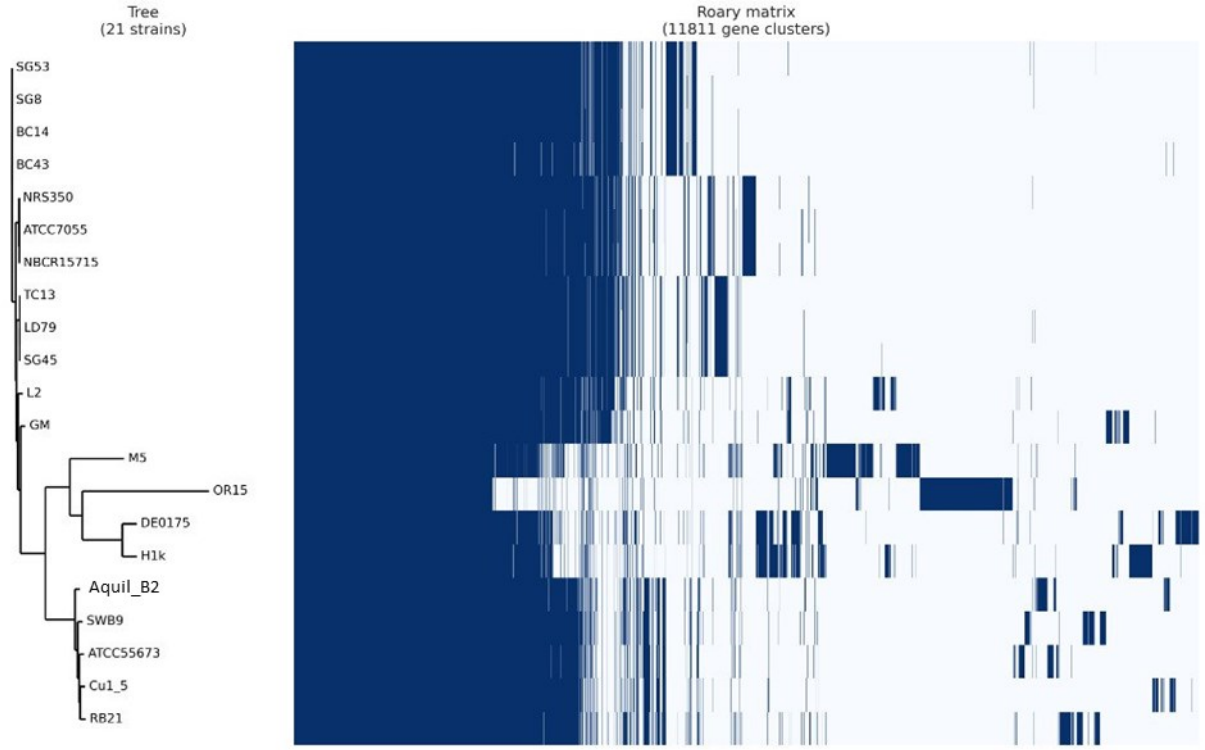


Figure 5, Roary matrix analysis of strain Aquil_B2

Strains Aquil_B3 and Aquil_B5, genetically close to each other (ANI = 99, 99, dDDH = 100, GC difference 0.01%), were characterized by a length of 5.06 Mbp and 5.07 Mbp, with a GC content of 41.26% and 41.27%, both with 6869 predicted genes. They clustered with the type strain *Peribacillus muralis* DSM 16288 with dDDH values of 81.60 and 81.50, an ANI value of 97.97 and a GC difference of 0.32%, confirming the identification also supported by the roary analysis (**Figure 6)**. It should be emphasized that the only other genome deposited as *P. muralis* is *P. muralis* strain G25-68 (assembly GCA_001646585.2), that for ANI and dDDH values is not identifiable as this species.

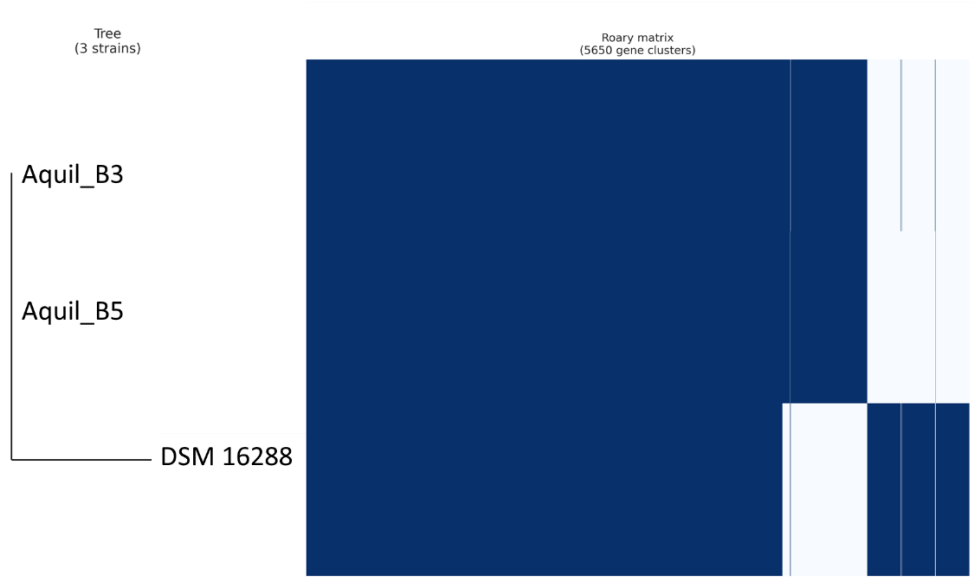


Figure 6, Roary matrix analysis of strain Aquil_B3 and Aquil_B5

Strain Aquil_B7, characterized by a length of 5.52 Mbp and 5445 genes, was classified as *Brevibacterium frigotolerans*, as it clustered with the reference strain FJAT-2396 (ANI = 98.52 dDDH = 86.10 GC difference 0.31%). Strain Aquil_B4, with a length of 6.68Mbp and 6727 genes also clustered with *B. frigotolerans* FJAT-2396 but having a higher phylogenetic distance with ANI of 96.42 and dDDH of 70.50, and especially with a %GC difference of 1.11%. The borderline dDDH value and the high difference in %GC therefore make the classification uncertain. The high heterogeneity of the strains deposited like this species can also be observed by the clustering of the genes performed by roary (**Figure 7)**. The *Peribacillus castrilensis* strain N3 which, as reported in **Supplementary table 1** showed values above the species separation threshold with both strains Aquil_B4 and Aquil_B7, was also included.

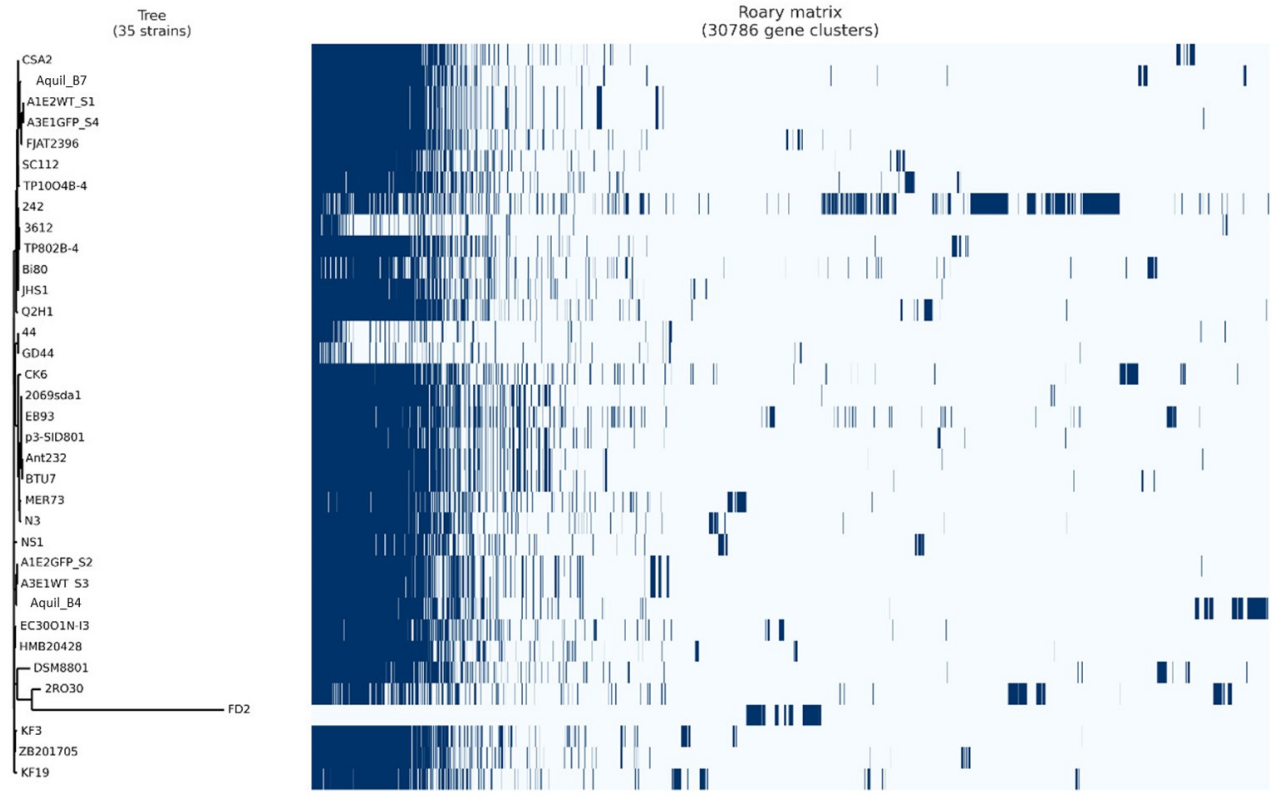


Figure 7, Roary matrix analysis of strain Aquil_B4 and Aquil_B7

Strains Aquil_B1 and Aquil_B8, both 5.65 Mbp in length with 5558 and 5557 genes respectively, appear to belong to the same species, as their comparison resulted in ANI values of 99.99 and dDDH of 100. They grouped with *Peribacillus simplex* NBRC 15720 reference strain but showing dDDH values of 65.6 and ANI of 93.50, resulting also close to *Brevibacterium frigotolerans* DSM 8801 with dDDH of 63.0 and ANI of 92.90. Comparing their genomes with all the genomes available on NCBI for this species, strains *P. simplex* I4 (ANI 95.60, dDDH 64.10) and *P. simplex* I6 (ANI 95.27, dDDH 63.30) resulted the most similar. Their classification is therefore uncertain, and as suggested by TYGS they could represent a new bacterial species. Strain Aquil_B6, characterized by a length of 4.26 Mbp and 4295 genes, was analyzed in detail in a previous work due to identification difficulties caused by the presence of incorrect reference sequences as reported in the previous chapter "Ancient Roman bacterium against current issues: strain Aquil_B6, *Paenisporosarcina quisquiliarum* or *Psychrobacillus psychrodurans*?". Briefly, after the TYGS analysis, it clustered with the "anomalous assemblies" GCA_900114885 and GCA_900109875 representing *Psychrobacillus psychrodurans* DSM 11713 (dDDH similarity of 88.6) and *Paenisporosarcina quisquiliarum* SK 55 (dDDH of 79.0). After resequencing of both genomes, it was possible to detect errors in the available *Paenisporosarcina quisquiliarum* SK 55 sequence, and to identify with certainty this isolate as *Psychrobacillus psychrodurans*.

Metabolisms

The metabolisms of the bacteria under examination were reconstructed by functional annotation performed using Prokka, classified by EggNOG, and mapped with KEGG. As reported in **Supplementary table 3**, all strains shared basic metabolic functions as glycolysis, gluconeogenesis, citrate cycle, ATP synthesis, fatty acid, and nucleotide metabolisms. Analysing in detail the main metabolic differences, regarding the carbohydrate metabolism, strains Aquil_B2 and Aquil_B6 were devoid of the pentose phosphate cycle, Aquil_B4 showed a complete degradation capacity of D-galacturonate and together with Aquil_B3, Aquil_B5 and Aquil_B6 also towards galactose, while Aquil_B2 resulted the only strain lacking in the biosynthetic capacity of glycogen. Evaluating the energy metabolism, Aquil_B6 did not show the capacity of assimilative reduction of sulphates and together with Aquil_B2 it also lacked the assimilative reduction of nitrates. In nucleotide metabolism, Aquil_B6 was the only strain without a pyrimidine degradation pathway. In the metabolism of amino acids, Aquil_B2 lacked both a synthetic pathway for the biosynthesis of cysteine and the methionine salvage pathway, while regarding the metabolism of lysine it lacked the pathways of succinyl and acetyl DAP for the synthesis of lysine, however it was the only strain who possessed a complete DAP dehydrogenase pathway. Considering the biosynthesis of polyamines, strain Aquil_B2 was the only one with a spermidine synthesis pathway starting from arginine. In the metabolism of cofactors and vitamins, compared to the others, strain Aquil_B6 did not have a metabolism for the biosynthesis of NAD starting from aspartate and the biosynthesis of biotin starting from pimeloyl ACP/CoA, and together with Aquil_B3 and Aquil_B5 it was not provided with the anaerobic metabolism for the biosynthesis of cobalamin. On the contrary, Aquil_B1, Aquil_B8, Aquil_B4 and Aquil_B7 showed the presence of a complete metabolic pathway for the biosynthesis of Pimeloyl-ACP. In the biosynthesis of terpenoids and polyketides and in xenobiotics degradation, strains Aquil_B1, Aquil_B8, Aquil_B4, Aquil_B7, Aquil_B3, and Aquil_B5 were found to possess metabolic pathways for the biosynthesis of isoprenoids C5 and C10-C20, while differently from the others, strain Aquil_B7 lacked the biosynthetic capacity of dTDP- L-rhamnose. Catechol meta-cleavage was not found in strains Aquil_B7, Aquil_B6 and Aquil_B2. As regards the drug resistance, a possible presence of methicillin resistance systems was suggested in strains Aquil_B1, Aquil_B8, Aquil_B4, and Aquil_B7, as well as a possible multidrug resistance efflux pump AbcA in all the strains genomes.

Bacteriocins, Prophages and CRISPR-Cas systems, Plasmids

Using Bagel4, the search for bacteriocins did not show the presence of strict hits with currently known bacteriocin sequences. Using PHASTER, the presence of prophages in genomic sequences was investigated (**Table 4**). A total of 3 intact and 2 questionable prophages were identified. Strains

Aquil_B1 and Aquil_B8 possessed the same prophages Phaenibacillus phage Tripp, and Bacillus phage PM1, while in strain Aquil_B2 the Paenibacillus phage Vegas was identified. In the current homologs of the identified species, the presence of prophages, absent in *B. muralis*, was found in only one strain of *B. frigoritolerans*, while in the case of *L. fusiformis*, 37 possible phages were found in 19 strains.

Table 4, prophages identified with PHASTER

Strain	Completeness	Phage	Accession number
Aquil_B1	intact	Bacillus phage PM1	NC_020883
	questionable	Paenibacillus phage Tripp	NC_028930
Aquil_B2	intact	Paenibacillus phage Vegas	NC_028767
Aquil_B3		-	
Aquil_B4		-	
Aquil_B5		-	
Aquil_B6		-	
Aquil_B7		-	
Aquil_B8	intact	Bacillus phage PM1	NC_020883
	questionable	Paenibacillus phage Tripp	NC_028930
L. fusiformis			
Cu1-5	intact	Bacillus phage SPP1	NC_004166
	intact	Bacillus phage phBC6A52	NC_004821
1226	questionable	Bacillus phage BM5	NC_029069
	questionable	Bacillus phage Gamma	NC_007458
	questionable	Brevibacillus phage Jenst	NC_028805
	questionable	Thermus phage phi OH2	NC_021784
S4C11	intact	Listeria phage B054	NC_009813
	questionable	Bacillus phage SPP1	NC_004166
	intact	Clostridium phage phiMMP03	NC_028959
RB-21	intact	Listeria phage B054	NC_009813
	intact	Paenibacillus phage Vegas	NC_028767
	intact	Paenibacillus phage Tadhana	NC_048691
ATCC 7055	intact	Lactobacillus phage Ld17	NC_025420
BC-43	intact	Paenibacillus phage Harrison	NC_028746
	questionable	Bacillus phage SPP1	NC_004166
G25-113	intact	Vibrio phage X29	NC_024369
GM	intact	Paenibacillus phage Harrison	NC_028746
	questionable	Bacillus phage SPP1	NC_004166
H1k	intact	Lactobacillus phage Ld3	NC_025421
L2	questionable	Listeria phage B054	NC_009813
	questionable	Bacillus phage SPP1	NC_004166
LD79	intact	Clostridium phage phiCD506	NC_028838
	intact	Bacillus phage vB_BhaS-171	NC_030904
M5	questionable	Clostridium phage CDMH1	NC_024144
	intact	Paenibacillus phage Vegas	NC_028767
NBRC 15717	intact	Lactobacillus phage Ld17	NC_025420
NEB1292	intact	Clostridium phage phiCD506	NC_028838
	intact	Thermus phage phi OH2	NC_021784
NRRL NRS-350	intact	Lactobacillus phage Ld17	NC_025420
SG8	intact	Paenibacillus phage Vegas	NC_028767
	questionable	Bacillus phage SPP1	NC_004166
SG45	intact	Bacillus virus 1	NC_009737
	intact	Clostridium phage phiCD506	NC_028838
SG53	intact	Paenibacillus phage Vegas	NC_028767
	questionable	Bacillus phage SPP1	NC_004166
TC-13	intact	Clostridium phage phiCTC2B	NC_030951
	intact	Bacillus phage vB_BhaS-171	NC_030904
B. frigoritolerans			

GD44	questionable	Aeribacillus phage AP45	NC_048651	

Concerning the CRISPR/Cas identified by CrisprCas Finder, only strain Aquil_B7 showed an IC-type Cas cluster consisting of 7 genes, and a 30-repetition CRISPR (**Table 5**), and also in this case 5 strains of *L. fusiformis* were found to possess different CRISPR/Cas systems, not found in the references of *B. muralis* and *B. frigoritolerans*.

Table 5, CRISPR/Cas systems

	Strain	Element	Spacer/Gene	Repeat consensus /cas genes	Evidence Level
	Aquil_B1	-	-	-	-
	Aquil_B2	-	-	-	-
	Aquil_B3	-	-	-	-
	Aquil_B4	-	-	-	-
	Aquil_B5	-	-	-	-
	Aquil_B6	-	-	-	-
	Aquil_B7	CRISPR	30	ATTTCAATCCACGCATCCATGAAGGATGCGAC	4
		CAS-TypeIC	7	Cas2_0_I-II-III-V, Cas1_0_IC, Cas4_0_I-II, Cas7_0_IC, Cas8c_0_IC, Cas5_0_IC, Cas3_0_I	
	Aquil_B8	-	-	-	-
L. fusiformis	1226	CRISPR	29	GTCTCTTCTCGTATGAGGAGAGTGGATTGAAAT	4
		CAS-TypeIC	7	Cas3_0_I, Cas5_0_IC, Cas8c_0_IC, Cas7_0_IC, Cas4_0_I-II, Cas1_0_IC, Cas2_0_I-II-III-V	
		CRISPR	31	GTCACTCCCTTTATGGGGAGTGTGGATTGAAAT	4
	BC-43	CAS-TypeIB	6	Cas6_0_I-III, Cas8a1_0_IB, Cas7_1_IB, Cas5_0_IB, Cas3_0_I, Cas2_0_I-II-III-V	
		CRISPR	9	TTCATCAACTAAGTGGAATGTGAA	4
	NEB1292	CRISPR	25	ATTTAAATTCCACTTAGTTAATGAAAAAC	4
		CAS-TypeIB	7	Cas6_0_I-III, Cas7_1_IB, Cas5_0_IB, Cas3_0_I, Cas4_0_I-II, Cas1_0_I-II-III, Cas2_0_I-II-III-V	
		CRISPR	28	GTTTTACATTAACTAAGTGGAATATAAAT	4
	SG8	CAS-TypeIB	6	Cas6_0_I-III, Cas8a1_0_IB, Cas7_1_IB, Cas5_0_IB, Cas3_0_I, Cas2_0_I-II-III-V	
		CRISPR	9	TTCATCAACTAAGTGGAATGTGAA	4
	SG53	CAS-TypeIB	6	Cas6_0_I-III, Cas8a1_0_IB, Cas7_1_IB, Cas5_0_IB, Cas3_0_I, Cas2_0_I-II-III-V	
		CRISPR	9	TTCATCAACTAAGTGGAATGTGAA	4

From the analysis conducted with Recycler, and subsequent blast alignment of the proposed sequences, 3 possible plasmids or contigs of plasmid origin were identified in strains Aquil_B3 (length 7263 bp), Aquil_B4 (length 19547 bp) and Aquil_B7 (length 8332 bp). As further evidence for the classification of the plasmidic contig of strain Aquil_B4, a plasmid recombination protein was identified (**Table 6**).

Table 6, length of predicted plasmids

Strain	Lenght of plasmid (bp)
Aquil_B1	-
Aquil_B2	-
Aquil_B3	7263
Aquil_B4	19547
Aquil_B5	8332
Aquil_B6	-
Aquil_B7	-
Aquil_B8	-

Resistance Genes

The presence of possible resistance genes was investigated using CARDS and ResFinder databases. No match with any reference from the database ResFinder was found, while through CARDS, the possible presence of the resistance gene vanRF in strain Aquil_B4 (identity> 94%), and the resistance gene dfrG in strain Aquil_B7 (% identity> 87%) were predicted. Together with these, **Table 7** shows the resistance genes identified in the strains deposited on NCBI with the use of the same databases, for the species with which the bacilli under analysis were identified. While in *B. muralis* and *L. fusiformis* sequences no evidence of possible resistance factors was found, *Brevibacterium frigotolerans* strains FJAT-2396, ZB201705 and GD44, shared with strain Aquil_B4 the possible presence of the vanRF gene. Upon closer analysis, the presence of an operon consisting of 3 genes emerged. The first gene was recognized as VanR-FM, with the function of response regulator transcription factor, the second as a possible histidine kinase, a function performed by the vanS gene, followed by the D-ala-D-ala carboxypeptidase of the VanY gene. By analyzing the functional annotation performed through PGAP, other possible factors of resistance to antibiotics in the genomes under analysis were also identified. In strains Aquil_B1, Aquil_B4, Aquil_B7 and Aquil_B8, the possible presence of AlzC and AlzD genes linked to resistance to azaleucine was highlighted, while in the same strains, together with the strain Aquil_B6, the possible presence of the FosM gene capable of conferring resistance to fosfomycin was identified. In strain Aquil_B4 the presence of a possible resistance factor to arsinothricin in an operon associated with arsenic resistance, adjacent to copper resistance factors was also highlighted.
Since many *L. fusiformis* have been investigated for their ability to resist toxic metals compounds, the presence of resistance genes to these compounds was also searched through the BacMet database (**Table 8**). On the basis of a sequence match rate> 79%, the widespread presence of two genes for aluminum resistance (ALU1-P and G2alt) and an arsenate reductase (arsC) was suggested, while only in strain Aquil_B2 it was possible to notice the presence of a quinone reductase (chrR).

Table 7, resistance genes identified using CARDS database

Species	Strain	RGI Criteria	ARO Term	AMR Gene Family	Drug Class	Resistance mechanism	% identity of matching region	% Lenght of reference sequence
B. frigoritolerans	**Aquil_B4**	Strict	vanRF	glycopeptide resistance gene cluster, vanR	glycopeptide antibiotic	antibiotic target alteration	94.37	100.00
	Aquil_B7	Strict	dfrG	trimethoprim resistant dihydrofolate reductase dfr	diaminopyrimidine antibiotic	antibiotic target replacement	87.27	100.00
	FJAT-2396	Strict	vanRF	glycopeptide resistance gene cluster, vanR	glycopeptide antibiotic	antibiotic target alteration	93.94	100.00
	ZB201705	Strict	vanRF	glycopeptide resistance gene cluster, vanR	glycopeptide antibiotic	antibiotic target alteration	95.24	100.00
	GD44	Strict	vanRF	glycopeptide resistance gene cluster, vanR	glycopeptide antibiotic	antibiotic target alteration	95.24	100.00
	Absent in : 3612, 242, 44							
B. muralis	DSM16288	Strict	mphM	macrolide phosphotransferase (MPH)	macrolide antibiotic	antibiotic inactivation	79.93	101.00
L. fusiformis	S4C11	Strict	ANT(4')-lb	ANT(4')	Aminoglycoside antibiotic	Antibiotic inactivation	96.84	100.00
		Strict	tet(45)	major facilitator superfamily (MFS) antibiotic efflux pump	Tetracycline antibiotic	Antibiotic efflux	85.59	100.00
	M5	Strict	clbA	Cfr 23S ribosomal RNA methyltransferase	lincosamide antibiotic, streptogramin antibiotic, oxazolidinone antibiotic, phenicol antibiotic, pleuromutilin antibiotic	Antibiotic target alteration	82.81	100.57
	DSM16288	Strict	mphM	macrolide phosphotransferase (MPH)	macrolide antibiotic	antibiotic inactivation	79.93	101.00
	Absent in : TC13, SWB9, SG53, SG45, SG8, RB21, NRS350, NBRC15717, LD79, L2, H1k, GM, DE0175, Cu1_5, BC43,BC14, ATCC55673, ATCC7055							

Table 8, metal resistance genes identified using BacMet database

Accession number	Gene		Strain							
			Aquil_B1	Aquil_B2	Aquil_B3	Aquil_B4	Aquil_B5	Aquil_B6	Aquil_B7	Aquil_B8
>BAC0489	ALU1-P	-	x	x	x	x	x	x	x	x
>BAC0581	arsC	Arsenate reductase	x	x	x	x	x	x	x	x
>BAC0490	G2alt	-	x			x			x	x
>BAC0539	chrR	Quinone reductase		x						
>BAC0003	acn	Aconitate hydratase	x	x	x	x	x	x	x	x
>BAC0012	actP	Copper-transporting P-type ATPase	x	x	x	x	x	x	x	x
>BAC0078	copA	Copper-exporting P-type ATPase	x	x	x	x	x	x	x	x
>BAC0079	copB	Copper-exporting P-type ATPase						x		
>BAC0101	ctpV	Probable copper-exporting P-type ATPase	x	x				x		x
>BAC0102	cueA	Copper transporter	x	x	x	x	x	x	x	x
>BAC0133	dnaK	Chaperone protein	x	x	x	x	x	x	x	x
>BAC0183	dnaK	Chaperone protein	x	x	x	x	x	x	x	x
>BAC0574	arsB	Arsenical pump membrane protein						x		
>BAC0575	arsB	Arsenical pump membrane protein						x		
>BAC0620	copA	Probable copper-importing P-type ATPase A	x	x	x	x	x	x	x	x
>BAC0622	copA	Probable copper-exporting P-type ATPase A	x	x		x		x	x	x
>BAC0629	copB	Copper-exporting P-type ATPase B						x		

Given the high presence of matches with high bitscore (evalue=0 and bit score > 570), but low match related to copper resistance factors (in addition to aconitate hydratase, chaperone proteins, arsenical pumps), an *in vivo* growth test was then performed to evaluate the resistance to copper. The results of the phenotypic characterization regarding Cu resistance performed on agar plates (**Table 9a**), showed that *B. muralis* strains Aquil_B3 and Aquil_B5 were inhibited in growth starting from 0.5 mM of $CuSO_4$, while strain Aquil_B7 was able to tolerate 0.5 mM, strains Aquil_B1, Aquil_B8, Aquil_B4 and Aquil_B6 to tolerate 1mM, up to the strain Aquil_B2 which showed growth up to 2mM levels of $CuSO_4$. On the contrary, in BHI broth test, all the strains were able to grow after 24 hours at a concentration of $CuSO_4$ up to 2,5 mM (**Table 9b**).

Table 9a, bacterial growth on Copper added BHI agar plates

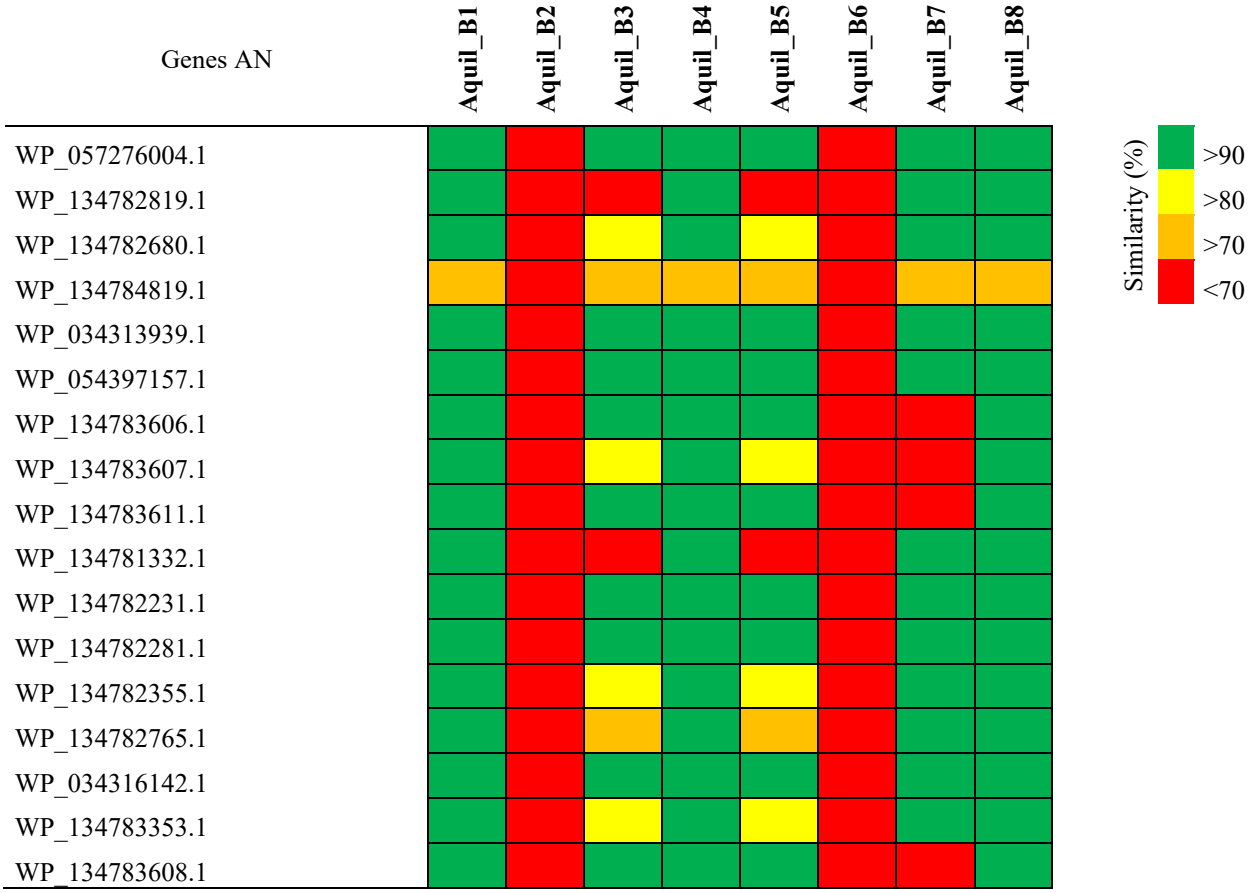
		$CuSO_4$ (mM)								
		0,5	**1**	**2**	**4**	**8**	**12**	**16**	**20**	**32**
24 hours	Aquil_B1	+	+	-	-	-	-	-	-	-
	Aquil_B2	+	+	+	-	-	-	-	-	-
	Aquil_B3	-	-	-	-	-	-	-	-	-
	Aquil_B4	+	-	-	-	-	-	-	-	-
	Aquil_B5	-	-	-	-	-	-	-	-	-
	Aquil_B6	+	+	-	-	-	-	-	-	-
	Aquil_B7	+	-	-	-	-	-	-	-	-
	Aquil_B8	+	+	-	-	-	-	-	-	-
48 hours	Aquil_B1	+	+	-	-	-	-	-	-	-
	Aquil_B2	+	+	+	-	-	-	-	-	-
	Aquil_B3	-	-	-	-	-	-	-	-	-
	Aquil_B4	+	+	-	-	-	-	-	-	-
	Aquil_B5	-	-	-	-	-	-	-	-	-
	Aquil_B6	+	+	-	-	-	-	-	-	-
	Aquil_B7	+	-	-	-	-	-	-	-	-
	Aquil_B8	+	+	-	-	-	-	-	-	-

Table 9b, mean values absorbance variation measured on copper added BHI broth

		$CuSO_4$ (mM)					
		0	**2,5**	**5**	**10**	**20**	**40**
24 hours	Aquil_B1	0,095	0,014	0,000	0,000	0,000	0,000
	Aquil_B2	0,480	0,031	0,000	0,000	0,000	0,000
	Aquil_B3	0,270	0,040	0,000	0,000	0,000	0,000
	Aquil_B4	0,132	0,038	0,000	0,000	0,000	0,000
	Aquil_B5	0,314	0,023	0,000	0,000	0,000	0,000
	Aquil_B6	0,155	0,019	0,000	0,000	0,000	0,000
	Aquil_B7	0,213	0,038	0,000	0,000	0,000	0,000
	Aquil_B8	0,209	0,034	0,000	0,000	0,000	0,000
48 hours	Aquil_B1	0,332	0,098	0,000	0,000	0,000	0,000
	Aquil_B2	0,628	0,113	0,000	0,000	0,000	0,000
	Aquil_B3	0,371	0,094	0,000	0,000	0,000	0,000
	Aquil_B4	0,190	0,077	0,000	0,000	0,000	0,000
	Aquil_B5	0,480	0,103	0,000	0,000	0,000	0,000
	Aquil_B6	0,511	0,100	0,000	0,000	0,000	0,000
	Aquil_B7	0,467	0,092	0,000	0,000	0,000	0,000
	Aquil_B8	0,297	0,091	0,000	0,000	0,000	0,000

The evaluation of the degradation capacity of phosphorus-based compounds was found in all the strains with the exception of Aquil_B2 and Aquil_B6, that did not show the presence of any gene linked to this function (**Table 10)**. In particular, strains Aquil_B1, Aquil_B8, and Aquil_B4 exhibited all the genes with a high % similarity (> 90%), with the exception of WP_134784819 gene, which was present with a lower similarity (> 70%).

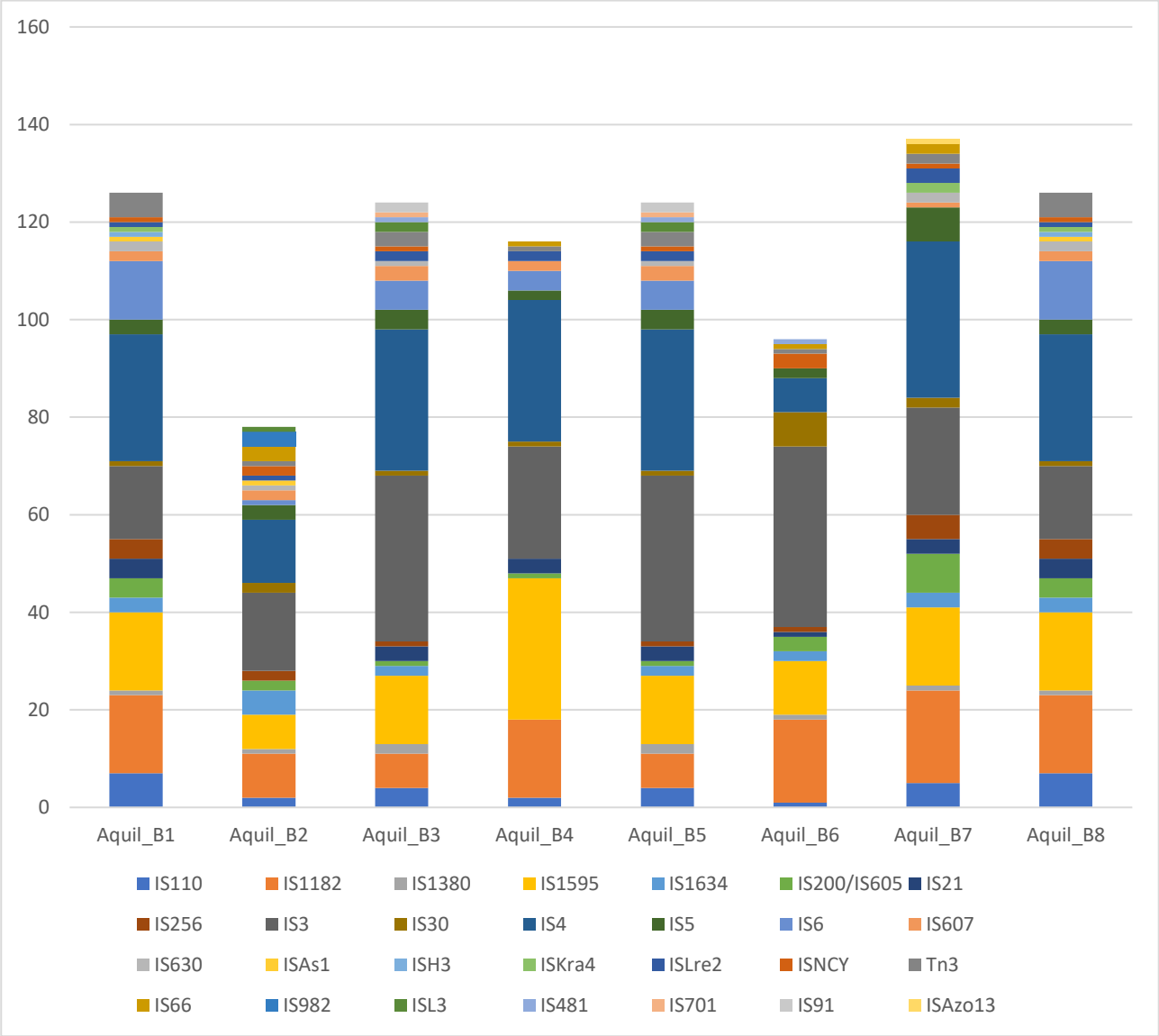
Table 10, phosphorous-based compounds degradation genes presence



Insertion sequences

The presence of possible IS sequences in the different genomes was analyzed using ISfinder (**Figure 8**). It can be noted that on average the possible IS most present are those of the IS3 and IS4 families, followed by IS1595 IS1182, IS6 and IS110.

Figure 8, IS sequences identified in the analyzed strains



For the strains in which the taxonomy was correctly identified, a comparison was made with the other genomes deposited for that species for the presence of IS (**Figure 9**). Considering *B. frigotolerans* specie, Aquil_B4 was characterized by a lower presence of IS than strain Aquil_B7 and by the absence of some groups of IS. However, when compared with other strains belonging to this species, they showed a high content of possible IS, lower only than in ZB201705 strain. On the contrary, *L. fusiformis* Aquil_B2 appeared to have less possible IS than the other strains deposited for this species. Strains Aquil_B3 and Aquil_B5 compared to the only one reference deposited for *P. muralis* (DSM16288), showed an equal number of possible IS sequences, but characterized by a greater number of IS4.

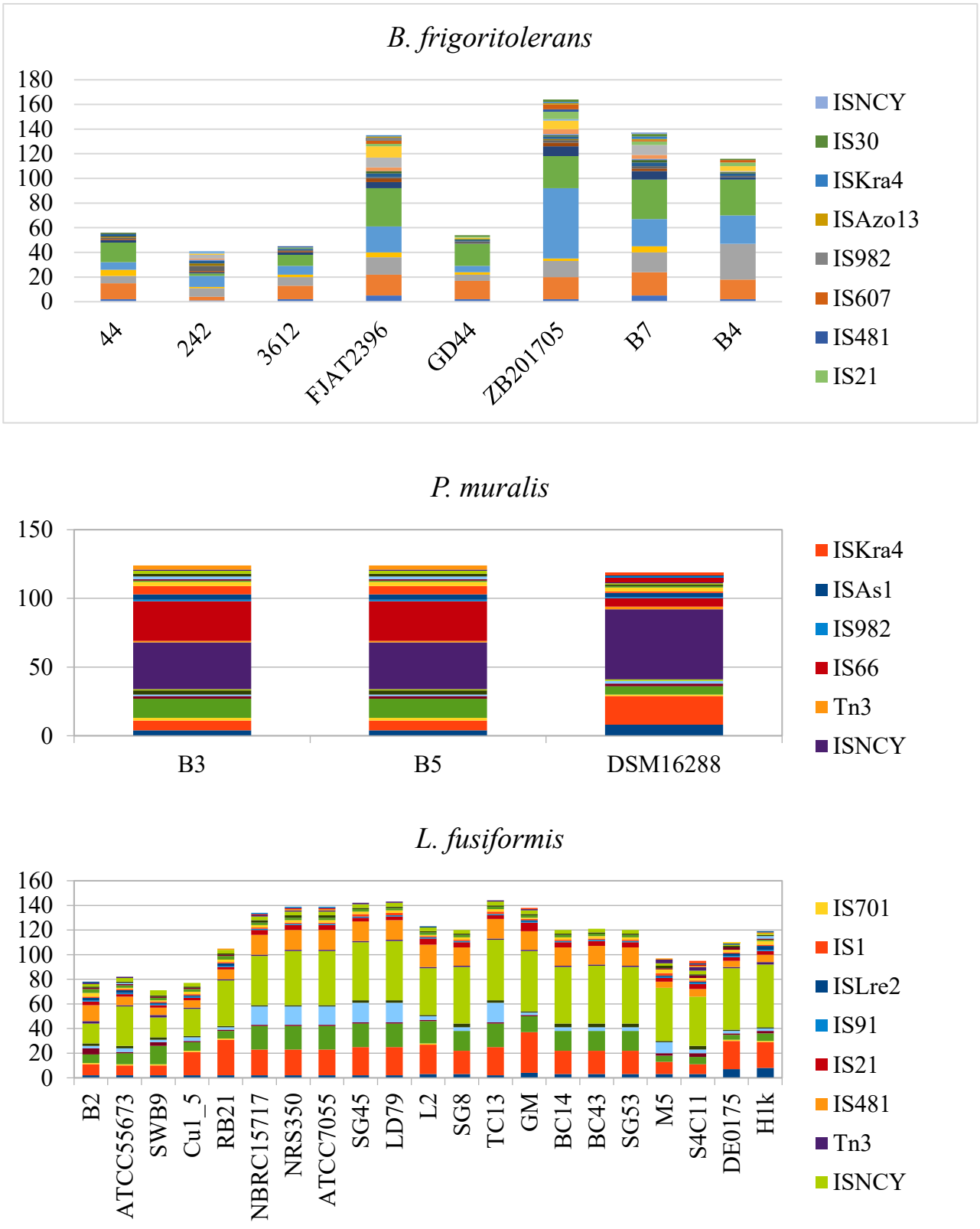


Figure 9, distribution of IS among the different species

8.4 Discussion and Conclusions

The microbial load found inside the amphora was not compatible with a contamination resulting from the sampling procedure. Furthermore, the exclusive finding of spore-forming species belonging to the *Bacillus* genus, and the absence of other microbial forms such as mold spores excluded possible contamination posthumous to the amphora sealing. We can therefore assume that the bacteria present inside derived exclusively from the initial contamination of the content, whose dehydration over time could be responsible for the high concentration of spores found. The 8 selected strains, as isolated from the count plate with highest dilution, are then probably one of the most present species inside the amphora and therefore the most representative. The 16S rRNA gene sequencing alone provided indications on preliminary species dentification, however, as expected, it failed to provide a unique identification given the high similarity that characterizes the different species with which a possible match has been identified. With the WGS analysis of these 8 species it was possible to obtain a precise identification only in some cases. Strain Aquil_B2 was classifiable confidently as *L. fusiformis*, thanks to the values of genetic similarity with the other strains deposited for this species. Strains Aquil_B3 and Aquil_B5 were identified as *P. muralis*, of which currently only a correct sequence belonging to type strain DSM 16288 is available on NCBI. Correct identification of strain Aquil_B6 as *P. psychrodurans* was only possible after resequencing of the reference strain. This result underlines once again the importance of checking the data that is uploaded to public databases, in order to avoid the propagation of errors (20). In the case of *P. psychrodurans* and *P. muralis* strains, however, it was more difficult to deepen the details of the comparisons given the small number of genomes available for these species. Both strains Aquil_B4 and Aquil_B7 clustered with the *B. frigotolerans* type strain for ANI and dDDH values. However, important differences have been highlighted between strain Aquil_B4 and the other strains deposited in the literature for this species. Strain Aquil_B4 showed a GC content of 39.51%, value that for the other deposited *B. frigotolerans* strains ranged between 40.60% (DSM 8801, FJAT-2396, 3621 strains) and 40.19% (CK6 strain), moreover with values that differ by more than 1% compared to the reference strain. In addition, strain Aquil_B4 showed the largest genome size, with a value of 6.68 Mb. The other strains deposited for this species in fact showed a smaller length between 4.98 Mb of the strain 242 and 6.29 Mb of the strain CK6. The difference in the %GC could therefore be linked to the greater length of the DNA sequence, which could be due to the lack of deletional bias, thus indicating possible evolutionary differences compared to the other deposited strains (21), however in contrast to what is reported in the literature where it is shown that larger genomes are associated with a greater G+C content (22). The pangenome analysis performed by Roary, underlines even more the high genetic heterogeneity of the strains identified for this species. It is in fact possible to note *B. frigotolerans* strains such as

242, 3612, GD44 and FD2, which differ massively in genetic content from the other strains. On the contrary, *Peribacillus castrilensis* strain N3, recently reported in the literature as a separate species (23), shares a high similarity with numerous other strains of *B. frigotolerans*, and from the taxonomy check performed by NCBI (ASM2101285v1) the strain shows an ANI similarity of 97.26% with the previous type strain *P. frigoritolerans* FJAT-2396 (24). Therefore, despite the recent reclassification of this species (25) from the evidence obtained, it does not seem that enough clarity has yet been cast on the genetic boundaries of this species. Further genetic studies on this species could be useful to clarify these aspects and deepen the knowledge on this bacterial species. Strains Aquil_B1 and Aquil_B8 clustered with *P. simplex* group, however ANI and dDDH values were lower than the threshold limit to classify these organisms. In fact, compared with the reference strain of the most similar species identified by TYGS, they showed values of dDDH <70% and ANI <95%, thresholds considered by the literature as a limit for the identification of a species (26)(27)(28)(29)(30). Values below the threshold were also found by comparing them with all the *P. simplex* strains available in the literature, further confirming their non-belonging to this species. A match with ANI values >95% was found with *P. simplex* strains I4 and I6, with which however the dDDH values were lower than 64%. These results suggest the usefulness of combining these two parameters to obtain better identification results. Further confirmation of the lack of match with known species for these two strains can also be seen in the Assembly QA tab on NCBI (ASM2260327v1 and ASM2260305v1). This function compares the deposited genomes with all available type-strain genomes to determine using ANI values the best matching reference. However, also in this case the best matching type-strain was found to be *Peribacillus simplex* NBRC 15720 (corresponded to DSM 1321), with ANI similarity values of 93.94, confirming the impossibility of identifying these strains with known species. Therefore, these two strains could be considered as a new identified species not yet sequenced. Regarding microbial resistance, the possible identification trough CARDS database of the vanRF gene in the genome of strain Aquil_B4, a feature also shared with current *B. frigotolerans* strains FJAT2396, ZB201705 and GD44, indicates that this type of antibiotic target alteration resistance gene was already present in this microbial group even before the extensive use of antibiotics. The identified operon can be identified with the typical structure of the VanRS two-component signal transduction system, consisting of the response regulator VanR and the sensor Kinase VanS, followed by the D-ala-D carboxypeptidase of the vanY gene (31)(32). Studies on the origin of this form of resistance were conducted on soil dwelling bacteria, identifying the presence of similar genes in different bacteria related to the bacilli under analysis (33). A further element of resistance characterizing *B. frigoritolerans* in the case of the strain Aquil_B7 is the possible presence of the dfrG gene, however not found in current organisms. The other strains Aquil_B3 Aquil_B5 and

Aquil_B2 belonging to *B. muralis* and *L. fusiformis* did not show the possible presence of resistance genes. Another possible resistance factor identified in strain Aquil_B4 is that against arsinothricin, a recently discovered antibiotic produced by soil bacteria (34). As reported in the literature, this factor was found in an operon associated with arsenic resistance, in the vicinity of another linked to copper resistance, indicating also in this case the presence in these bacteria whose ecological niche is linked to the soil of resistance factors to these heavy metals. The study of resistance capacities and degradation of toxic compounds by microorganisms related to the genus *Bacillus* spp. is particularly important, as it may allow their potential use in the improvement of soil quality in the future (35). In this regard, in this study different factors have emerged that confer resistance also to other metals reported in several articles for different bacilli (15)(16)(17)(18)(19), in this case both *in silico* and *in vitro*. In fact, several possible copper resistance genes have been identified, as confirmed *in vitro* by the resistance test to $CuSO_4$, where all strains tolerated a concentration of this substance up to 2,5 mM as observed in the BHI broth test, values in agreement with other studies, but lower than some particularly resistant strains of the genus *Bacillus spp*. (36)(37). Another capacity found, with the exception of strains Aquil_B2 and Aquil_B6, is the presence of a series of genes, reported in the literature as functional to the degradation of phosphorus-based compounds. As regards the presence of prophages, their presence has been found in strains Aquil_B1, Aquil_B8 and Aquil_B2. It should be emphasized that one factor that seems to characterize the *L. fusiformis* group is the high presence of prophages, found both in our Aquil_B2 strain with 2 prophages identified as intact and 1 as questionable, and in the reference genomes, where were found in almost all strains, with 37 possible prophages identified. This feature can also be linked to the greater presence of possible CRISPR/Cas phage defence systems, found in 5 reference strains, of which 4 CAS type IB systems and 1 type IC system, however not found in the strain under examination *L. fusiformis* Aquil_B2. It was however possible to identify an IC type system in the genome of strain Aquil_B7.

8.5 Materials and Methods

Sampling

For transfer from the warehouse of the Archaeological Museum of Aquileia to the Microbiology Laboratory at the Department of Agri-food, Environmental and Animal Sciences of the University of Udine the amphora was carefully packed with Pluriball plastic material (110g/m^2, Packing-2000), after having externally strengthened the structure with gauze. Upon arrival, the amphora was intact and sealed with cementitious material. It was opened under a laminar flow hood in the presence of a Bunsen burner after flaming the outside. The employed instruments were sterilized in an autoclave at 121 °C for 15 min or when incompatible with humidity, sterilization was carried out by air in an oven at 150 °C for 12 hours. GMP were followed to avoid contaminations. The organic material present inside the amphora (Figure XY) was collected under sterile condition and placed inside sterile stomacher bags, which were immediately placed under vacuum to avoid oxidation of the material and stored at 5 °C pending analysis.

Microbiological analysis

The microbiological analyses were carried out on the material contained in the ancient Roman amphora. Ten grams of the contents were withdrawn in a sterile stomacher bag and diluted 1:10 with saline-peptone water (8 g/L NaCl, 1 g/L bacteriological peptone; Oxoid, Italy, distilled water 1000 mL) and mixed for 2.5 min in a Stomacher machine (PBI, Milan, Italy). Further decimal dilutions were made in the same solution and the following microbiological analyses were performed in duplicate agar plates on three biological replicates. Ten (10) different culture media were used for standard plate counts: Plate Count Agar (PCA, Oxoid, Italy) and Brain heart infusion agar (BHI, Oxoid, Italy) incubated at 10, 30 and 42 °C under aerobic conditions were used to investigate the total psychrophilic, mesophilic and thermophilic microbial counts respectively; De Man, Rogosa and Sharpe (MRS, Oxoid, Italy) incubated at 30 °C under aerobic and microaerophilic conditions was used to evaluate the presence of lactic acid bacteria or other bacteria nutritionally demanding; Sulphite Polymyxin Sulphadiazine (SPS, Oxoid, Italy) and Differential Reinforced Clostridial Medium (DRCM, Sigma-Aldrich, Italy) were used to evaluate the presence of anaerobic spore-forming bacteria; Yeast Extract Peptone Dextrose (YPD, Sigma-Aldrich, Italy), Wallerstein Laboratory (WL) (Oxoid, Italy) and Rose Bengal Agar (RBA, BioLife Italiana, Italy) incubated at 25 °C under aerobic conditions were used to look for yeast and moulds; Mannitol Salt agar incubated at 30 °C under aerobic conditions (MSA, Oxoid, Italy) was used to evaluate the presence of *Staphylococcus* spp. and micrococci; Vibrio ChromoSelect Agar (VCS, Sigma-Aldrich, Italy) incubated at 35 °C for 24-48 h was used for the enumeration of *Vibrio* spp. For *Listeria*

monocytogenes (ISO/DIS11290-1 method, 1990) and *Salmonella* spp. (ISO/DIS 6579 methods, 1991) the International Organization for Standardization ISO/DIS methods were performed. After counting means and standard deviations were calculated. When possible, an average of 10 colonies were randomly selected from the agar plates. Colonies were selected independently from their morphology, color or size. Using the plate counting method, taking an aliquot of 0.1 mL or 1 mL depending on the type of medium, the bacterial count of the different species possibly present, using the incubation protocol (temperature, oxygen percentage and time) specified by the manufacturer, was carried out using the following mediums: Plate Count Agar (PCA), Brain Heart Infusion (BHI), De Man, Rogosa and Sharpe (MRS), Sulphite Polymyxin Sulphadiazine (SPS), Differential Reinforced Clostridial Medium (DRCM), Yeast Extract Peptone Dextrose (YPD), Wallerstein Laboratory (WL) (Oxoid, Italy) . The bacterial count was carried out on the plates containing between 30 and 150 colonies, while the isolation of the strains to be analyzed was performed by sampling all the colonies present on the count plate with the highest dilution to minimize the possibility of considering bacteria deriving from possible environmental contamination during the sampling of the amphora.

Molecular and phenotypic characterization of isolates

Isolates were subjected to Gram staining, catalase, and peroxidase tests. In order to eliminate clones, molecular characterization was performed comparing the genetic fingerprints obtained by RAPD, Rep-PCR and SAU-PCR techniques (38). Briefly, DNA extraction for PCR-based genetic fingerprint of the strains was performed using the GenElute™ Bacterial Genomic DNA Kit (Sigma-Aldrich, Milan, Italy) following the manufacturer instruction. DNA was then quantified and standardized at 50 ng/µL (Nanodrop One system, Thermo Scientific, Marietta, OH, USA). PCRs were performed using the amplification condition described by Iacumin et al. (2006) (38) using a C1000 Touch Thermal Cycler (BioRad, Milan, Italy). The following primers were used for the reactions: M13 (5'-GAG GGT GGC GGT TCT-30'), (GTG) 5 (5'-GTGGTGGTGGTGGTG-3') and SAG1 (5'-CCGCCGCGATCAG-3') for RAPD, Rep-PCR and SAU-PCR, respectively. Electrophoresis was performed on 1.5% (w/v) agarose gels in 0.5X TBE at 120 V for 6 h. Gels were externally stained by incubation in 0.5X TBE buffer containing 0.5 µg/ml ethidium bromide (Sigma-Aldrich, Milan, Italy) for 30 min in the dark. Pictures of the gels were digitally captured using the BioImaging System GeneGenius (SynGene) and the pattern analysis software package Gel Compare Version 4.1 (Applied Maths, Kortrijk, Belgium) was used for the analysis. Similarities in the profiles of bands were evaluated based on Pearson product-moment correlation coefficients. Dendrograms were obtained by the UPGMA clustering algorithms (39). The assessment of the sporulation capacity was carried out by staining with Malachite Green. From overnight culture streaked on BHI of each strain was

cultivated on Nutrient Agar and incubated at 37 ° C for 5 days, then fixed on a slide and stained with aqueous solution of Malachite Green (Certistain, Merck Life Science Srl, Milan, Italy) at 10% and observed under the microscope. The $CuSO_4$ resistance test was performed following the protocol suggested by Cai et al. (2019) (40) for broth culture and Glibota et al. (2019) (41) for plate growth. In the first method, 10 µL of the overnight culture of each strain was inoculated in 5 mL of sterile BHI broth at increasing concentrations of $CuSO_4$ equal to 2.5, 5, 10, 20 and 40 mM/L. The tubes were then incubated at 30 ° C and their optical density was measured at 600 nm with the Bio-Rad SmartSpec3000 spectrophotometer (BioRad, Italy) at 24 and 48 hours. In the second method 20 mL plates of BHI Agar added with $CuSO_4$ were prepared at the following concentrations: 0.5, 1, 2, 4, 8, 12, 16, 20, 32 mM/L. For each of these media, 5 µL overnight culture spots were inoculated in BHI broth from each strain. The plates were then incubated at 30 ° C and checked at 24 and 48 hours.

Genome sequencing, annotation, and characterization

The genomes were sequenced and assembled as reported in Colautti et al. (2022) (42). The amplification and sequencing (Sanger technology, Eurofins Genomics, Germany) of V1-V3 region of 16S rDNA was performed to preliminary identify the strains. This procedure was carried out using P1 and P4 primers (43), with the following amplification conditions: final volume 50 µL, 10 mM Tris–HCl, pH 8, KCl 50 mM, $MgCl_2$ 1.5 mM, dNTPs 0.2 mM, each primer 0.2 µM, 1.25 U *Taq*-polymerase (Applied Biosystem, I) with 100 ng of DNA. For the whole sequencing process, each strain was cultured in BHI broth (Oxoid, Italy). From the cell pellet obtained from centrifugation for 5 min at 5000 × *g*, the bacterial DNA was extracted via MagAttract HMW DNA kit (Qiagen, Germany) and fragmented by sonication via BioRuptor (Diagenode, Belgium), and libraries were prepared with Celero DNA-Seq kit (Tecan, Swiss). Libraries were then sequenced in paired end mode to obtain reads with a length of 300 bp with the MiSeq platform (Illumina, USA). The reads thus obtained were then carefully assembled via the WGA-LP pipeline (44) using SPAdes as the assembler (45), verifying the quality of the initial reads and the assembly obtained through CheckM (46) and Quast (47). The taxonomic identification of the obtained genomes was carried out through the Type Strain Genome Server and the calculation of the dDDH through the Genome to Genome Distance Calculator using formula 2 as a reference (48), while fastANI (29) was used for ANI analysis. The genomes were annotated using Prokka (49) and PGAP (50), and analyzed for their genetic characteristics with Roary (51). The KEGG annotation was carried out through EggNOG (52), and the metabolic pathways were obtained with the KEGG Mapper function from KEGG (53). Phages were searched with Phaster (54) and CRISPR/Cas with CRISPRCasFinder (55), while possible plasmids or contigs of possible plasmidic origin were identified with Recycler (56) and aligning the suggested contigs to the BLAST suite. The ISfinder database was used to search for Insertion

Sequences (IS) using the default settings (57). The accession number of the reference genomes with which the analyzes were carried out are shown in **Table 13**.

Virulence and Resistance factors

RGI from CARDS (58) and ResFinder (59) were used to search for resistance and virulence genes, while to search for antibacterial biocides and metal resistance genes, the BacMet database (version 1.0, experimentally confirmed dataset) (60) was used. The evaluation of the presence of genes related to the degradative capacity of phosphorus compounds was made by blasting the genes suggested by Jin et al. (2020) (19) .

8.5 References

1. Seaward MRD, Cross T, Unsworth BA. 1976. Viable bacterial spores recovered from an archaeological excavation. Nature 261:407–408.

2. Andam CP, Worby CJ, Chang Q, Campana MG. 2016. Microbial Genomics of Ancient Plagues and Outbreaks. Trends Microbiol 24:978–990.

3. Bos KI, Kühnert D, Herbig A, Esquivel-Gomez LR, Andrades Valtueña A, Barquera R, Giffin K, Kumar Lankapalli A, Nelson EA, Sabin S, Spyrou MA, Krause J. 2019. Paleomicrobiology: Diagnosis and evolution of ancient pathogens. Annu Rev Microbiol 73:639–666.

4. Lugli GA, Milani C, Mancabelli L, Turroni F, Ferrario C, Duranti S, Sinderen D van, Ventura M. 2017. Ancient bacteria of the ötzi's microbiome: A genomic tale from the Copper Age. Microbiome 5:1–18.

5. Cano RJ, Rivera-Perez J, Toranzos GA, Santiago-Rodriguez TM, Narganes-Storde YM, Chanlatte-Baik L, García-Roldán E, Bunkley-Williams L, Massey SE. 2014. Paleomicrobiology: Revealing fecal microbiomes of ancient indigenous cultures. PLoS One 9.

6. Dagli N, Dagli R, Baroudi K, Tarakji B. 2015. Oral Paleomicrobiology: Study of Ancient Oral Microbiome. J Contemp Dent Pract 16:588–594.

7. Nodari R, Drancourt M, Barbieri R. 2021. Paleomicrobiology of the human digestive tract: A review. Microb Pathog 157:19–21.

8. Darling MI, Donoghue HD. 2014. Insights from paleomicrobiology into the indigenous peoples of pre-colonial America - A Review. Mem Inst Oswaldo Cruz 109:131–139.

9. Setlow P. 2006. Spores of Bacillus subtilis: Their resistance to and killing by radiation, heat and chemicals. J Appl Microbiol 101:514–525.

10. Čavka M, Glasnović A, Janković I, Šikanjić PR, Perić B, Brkljačić B, Mlinarić-Missoni E, Škrlin J. 2010. Microbiological analysis of a mummy from the archeological museum in Zagreb. Coll Antropol 34:803–805.

11. Southern PM. 2008. Archeology meets clinical microbiology: Analysis of the microbial content of soil from archeological sites in Italy and Belize. Lab Med 39:601–602.

12. Wright GD, Poinar H. 2012. Antibiotic resistance is ancient: Implications for drug discovery. Trends Microbiol 20:157–159.

13. Afouda P, Dubourg G, Levasseur A, Fournier PE, Delerce J, Mediannikov O, Diene SM,

Nahon D, Bourlès D, Rolain JM, Raoult D. 2020. Culturing ancient bacteria carrying resistance genes from permafrost and comparative genomics with modern isolates. Microorganisms 8:1–13.

14. Cipriano MT, Carre MB. 1987. Note sulle anfore conservate nel museo di Aquileia. Antich Altoadriatiche XXIX 2:479–494.

15. He M, Li X, Liu H, Miller SJ, Wang G, Rensing C. 2011. Characterization and genomic analysis of a highly chromate resistant and reducing bacterial strain Lysinibacillus fusiformis ZC1. J Hazard Mater 185:682–688.

16. Ibeanusi V. 2018. "Genome-Centric Evaluation of Bacillus sp. strain –ATCC55673 and Response to Uranium Biomineralization." Significances Bioeng Biosci 2:157–164.

17. Cai X, Zheng X, Zhang D, Iqbal W, Liu C, Yang B, Zhao X, Lu X, Mao Y. 2019. Microbial characterization of heavy metal resistant bacterial strains isolated from an electroplating wastewater treatment plant. Ecotoxicol Environ Saf 181:472–480.

18. Jibrin AM, Oyewole OA, Yakubu JG, Hussaini A. 2020. Heavy metals biosorption by urease producing Lysinibacillus fusiformis 5B. Eur J Biol Res 10:326–335.

19. Jin M, Zhao Q, Zhou Z, Zhu L, Zhang Z, Jiang L. 2020. Draft Genome Sequence of a Potential Organic Phosphorus-Degrading Bacterium Brevibacterium frigoritolerans GD44, Isolated from Radioactive Soil in Xinjiang, China. Curr Microbiol 77:2896–2903.

20. Bagheri H, Severin AJ, Rajan H. 2020. Detecting and correcting misclassified sequences in the large-scale public databases. Bioinformatics 36:4699–4705.

21. Mira A, Ochman H, Moran NA. 2001. Deletional bias and the evolution of bacterial genomes. Trends Genet 17:589–596.

22. Almpanis A, Swain M, Gatherer D, McEwan N. 2018. Correlation between bacterial G+C content, genome size and the G+C content of associated plasmids and bacteriophages. Microb genomics 4:0–7.

23. Rodríguez M, Reina JC, Sampedro I, Llamas I, Martínez-Checa F. 2022. Peribacillus castrilensis sp. nov.: A Plant-Growth-Promoting and Biocontrol Species Isolated From a River Otter in Castril, Granada, Southern Spain. Front Plant Sci 13:1–14.

24. Zhang C, Li X, Yin L, Liu C, Zou H, Wu Z, Zhang Z. 2019. Analysis of the complete genome sequence of Brevibacterium frigoritolerans ZB201705 isolated from drought- and salt-stressed

rhizosphere soil of maize. Ann Microbiol 69:1489–1496.

25. Montecillo JA V., Bae H. 2022. Reclassification of Brevibacterium frigoritolerans as Peribacillus frigoritolerans comb. nov. based on phylogenomics and multiple molecular synapomorphies. Int J Syst Evol Microbiol 72:1–9.

26. Auch AF, von Jan M, Klenk HP, Göker M. 2010. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2:117–134.

27. Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A 106:19126–19131.

28. Ciufo S, Kannan S, Sharma S, Badretdin A, Clark K, Turner S, Brover S, Schoch CL, Kimchi A, DiCuccio M. 2018. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. Int J Syst Evol Microbiol 68:2386–2392.

29. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:1–8.

30. Wang Q, Zhu W, Wang E, Zhang L, Li X, Wang G. 2016. Genomic identification of rhizobia-related strains and threshold of ANI and core-genome for family, genus and species. Int J Environ Agric Res ISSN 2:76–86.

31. Hong H, Hutchings MI. 2008. Two-Component Systems.

32. Meziane-Cherif D, Stogios PJ, Evdokimova E, Savchenko A, Courvalin P. 2014. Structural basis for the evolution of vancomycin resistance D,D-peptidases. Proc Natl Acad Sci U S A 111:5872–5877.

33. Fraimow H, Knob C, Herrero IA, Patel R. 2005. Putative VanRS-like two-component regulatory system associated with the inducible glycopeptide resistance cluster of Paenibacillus popilliae. Antimicrob Agents Chemother 49:2625–2633.

34. Nadar VS, Chen J, Dheeman DS, Galván AE, Yoshinaga-Sakurai K, Kandavelu P, Sankaran B, Kuramata M, Ishikawa S, Rosen BP, Yoshinaga M. 2019. Arsinothricin, an arsenic-containing non-proteinogenic amino acid analog of glutamate, is a broad-spectrum antibiotic. Commun Biol 2.

35. Fakhar A, Gul B, Gurmani AR, Khan SM, Ali S, Sultan T, Chaudhary HJ, Rafique M, Rizwan

M. 2022. Heavy metal remediation and resistance mechanism of Aeromonas, Bacillus, and Pseudomonas: A review. Crit Rev Environ Sci Technol 52:1868–1914.

36. Üreyen Esertaş ÜZ, Uzunalioğlu E, Güzel Ş, Bozdeveci A, Alpay Karaoğlu Ş. 2020. Determination of bioremediation properties of soil-borne Bacillus sp. 5O5Y11 and its effect on the development of Zea mays in the presence of copper. Arch Microbiol 202:1817–1829.

37. He M, Li X, Guo L, Miller SJ, Rensing C, Wang G. 2010. Characterization and genomic analysis of chromate resistant and reducing Bacillus cereus strain SJ1. BMC Microbiol 10.

38. Iacumin L, Comi G, Cantoni C, Cocolin L. 2006. Molecular and technological characterization of Staphylococcus xylosus isolated from naturally fermented Italian sausages by RAPD, Rep-PCR and Sau-PCR analysis. Meat Sci 74:281–288.

39. Vauterin, L. & Vauterin P. 1992. Computer-aided objective comparison of electrophoresis patterns for grouping and identification of microorganisms. Eur Microbiol 1:37–42.

40. Cai Z, Du, Wenlong, Liu L, Pan D, Lu L. 2019. Molecular Characteristics of the Conserved Aspergillus nidulans. mSphere 4:1–11.

41. Glibota N, Grande Burgos MJ, Gálvez A, Ortega E. 2019. Copper tolerance and antibiotic resistance in soil bacteria from olive tree agricultural fields routinely treated with copper compounds. J Sci Food Agric 99:4677–4685.

42. Colautti A, Comi G, Paoli E De, Peterlunger E, Novello M, Braidotti E, Pasini D, Iacumin L. Draft Genome Sequences of Eight Bacilli Isolated from an Ancient Roman Amphora 39:11–13.

43. Klijn N, Weerkamp AH, De Vos WM. 1991. Identification of mesophilic lactic acid bacteria by using polymerase chain reaction-amplified variable regions of 16S rRNA and specific DNA probes. Appl Environ Microbiol 57:3390–3393.

44. Rossi N, Colautti A, Iacumin L, Piazza C. 2021. WGA-LP: a pipeline for whole genome assembly of contaminated reads. Bioinformatics 1–3.

45. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477.

46. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: Assessing

the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055.

47. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.

48. Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun 10.

49. Seemann T. 2014. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.

50. Zhao Y, Wu J, Yang J, Sun S, Xiao J, Yu J. 2012. PGAP: Pan-genomes analysis pipeline. Bioinformatics 28:416–418.

51. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693.

52. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Huerta-Cepas J. 2021. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. bioRxiv 5:2021.06.03.446934.

53. Kanehisa M, Sato Y. 2020. KEGG Mapper for inferring cellular functions from protein sequences. Protein Sci 29:28–35.

54. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–W21.

55. Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res 35:52–57.

56. Rozov R, Kav AB, Bogumil D, Shterzer N, Halperin E, Mizrahi I, Shamir R. 2017. Recycler: An algorithm for detecting plasmids from de novo assembly graphs. Bioinformatics 33:475–482.

57. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. 2006. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res 34:32–36.

58. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen AL V., Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran HK, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma

AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar G V., McArthur AG. 2020. CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 48:D517–D525.

59. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, Philippon A, Allesoe RL, Rebelo AR, Florensa AF, Fagelhauer L, Chakraborty T, Neumann B, Werner G, Bender JK, Stingl K, Nguyen M, Coppens J, Xavier BB, Malhotra-Kumar S, Westh H, Pinholt M, Anjum MF, Duggett NA, Kempf I, Nykäsenoja S, Olkkola S, Wieczorek K, Amaro A, Clemente L, Mossong J, Losch S, Ragimbeau C, Lund O, Aarestrup FM. 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother 75:3491–3500.

60. Pal C, Bengtsson-Palme J, Rensing C, Kristiansson E, Larsson DGJ. 2014. BacMet: Antibacterial biocide and metal resistance genes database. Nucleic Acids Res 42:737–743.

8.6 Supplementary material

Supplementary table 1, accession number of the strains used for comparisons

Species	Strain	WGS/Chromosome Accession Number
P. frigoritolerans	JHS1	GCA_022394675.1
P. frigoritolerans	CK6	GCA_023823915.1
P. frigoritolerans	Q2H1	GCA_025960625.1
P. frigoritolerans	KF3	GCA_023823955.1
P. frigoritolerans	Ant232	GCA_021537535.1
P. frigoritolerans	NS1	GCA_024758165.1
P. frigoritolerans	ZB201705	GCA_004006475.1
P. frigoritolerans	HMB20428	GCA_026013865.1
P. frigoritolerans	KF19	GCA_023823995.1
P. frigoritolerans	DSM 8801	GCA_024169475.1
P. frigoritolerans	FJAT-2396	GCA_001636405.1
P. frigoritolerans	2069sda1	GCA_024159205.1
P. frigoritolerans	2RO30	GCA_025209795.1
P. frigoritolerans	44	GCA_014876535.1
P. frigoritolerans	3612	GCA_007828935.1
P. frigoritolerans	GD44	GCA_004525735.1
P. frigoritolerans	Aquil_B7	GCA_022603155.1
P. frigoritolerans	MER 73	GCA_023714145.1
P. frigoritolerans	A1E2WT_S1	GCA_023502905.1
P. frigoritolerans	CSA2	GCA_025548515.1
P. frigoritolerans	A3E1GFP_S4	GCA_023502885.1
P. frigoritolerans	A1E2GFP_S2	GCA_023502945.1
P. frigoritolerans	A3E1WT_S3	GCA_023502825.1
P. frigoritolerans	SC112	GCA_024733505.1
P. frigoritolerans	BTU7	GCA_021728995.1
P. frigoritolerans	p3-SID801	GCA_025142885.1
P. frigoritolerans	Bi80	GCA_918698165.1
P. frigoritolerans	EB93	GCA_009996885.1
P. frigoritolerans	Aquil_B4	GCA_022603205.1
P. frigoritolerans	242	GCA_018195605.1
P. muralis	DSM 16288	GCA_001439925.1
P. muralis	G25-68	GCA_001645685.2
L. fusiformis	1226	GCA_007362955.1
L. fusiformis	ATCC 7055	GCA_003049525.1
L. fusiformis	ATCC 55673	GCA_008795865.1
L. fusiformis	BC-14	GCA_900104275.1
L. fusiformis	BC-43	GCA_900116155.1
L. fusiformis	Cu1-5	GCA_007923505.1
L. fusiformis	DE0175	GCA_007678325.1
L. fusiformis	G25-113	GCA_015845625.1
L. fusiformis	GM	GCA_002358065.1

L. fusiformis	H1k	GCA_000633275.1
L. fusiformis	Juneja	GCA_002845985.1
L. fusiformis	L2	GCA_013112215.1
L. fusiformis	LD79	GCA_900102355.1
L. fusiformis	M5	GCA_001726065.1
L. fusiformis	NBRC 15717	GCA_006540205.1
L. fusiformis	NEB1292	GCA_016925635.1
L. fusiformis	NRRL NRS-350	GCA_003367495.1
L. fusiformis	OR-15	GCA_016308375.1
L. fusiformis	RB-21	GCA_000724775.3
L. fusiformis	S4C11	GCA_015161405.1
L. fusiformis	SG8	GCA_900101005.1
L. fusiformis	SG45	GCA_900110175.1
L. fusiformis	SG53	GCA_900113785.1
L. fusiformis	SW-B9	GCA_000755455.1
L. fusiformis	TC-13	GCA_900110625.1
L. fusiformis	ZB2	GCA_000313955.2
L. fusiformis	ZC1	GCA_000178135.1
P. simplex	313	GCA_002287755.1
P. simplex	7894	GCA_002276345.1
P. simplex	BA2H3	GCA_000785385.1
P. simplex	BE23	GCA_003931555.1
P. simplex	CFBP13531	GCA_014841365.1
P. simplex	DE0003	GCA_007680885.1
P. simplex	DE0084	GCA_007679575.1
P. simplex	DSM 1321	GCA_002243645.1
P. simplex	GGC-P6A	GCA_007786515.1
P. simplex	I4	GCA_005217225.1
P. simplex	I6	GCA_005217145.1
P. simplex	MGYG-HGUT-00083	GCA_902363015.1
P. simplex	MYb48	GCA_002979275.1
P. simplex	NBRC 157020	GCA_001591785.1
P. simplex	OG2	GCA_002276655.1
P. simplex	P558	GCA_900000145.1
P. simplex	RUG2-6	GCA_900156045.1
P. simplex	SH-B26	GCA_001578185.1
P. simplex	VanAntwerpen02	GCA_001542915.1
P. simplex	WY10	GCA_002351505.1

Supplementary table 2, ANI and dDDH values comparison

Query	References	ANI	dDDH	G+C difference
B1	B8	100	100.00	0.00
	I4	95,59	64.10	0.14
	I6	95,25	63.30	0.20
	NBRC17520	93,5	52.90	0.44
	GGC-P6A	93,46	53.00	0.59
	DSM1321	93,41	53.10	0.35
	MGYG-HGUT-00083	93,37	52.90	0.64
	DE003	93,28	52.30	0.14
	7894	93,25	51.40	0.02
	DE0084	93,16	52.40	0.46
	CFBP13531	93,14	51.40	0.06
	P558	93,12	51.30	0.01
	MYb48	93,07	51.30	0.03
	OG2	93,04	51.10	0.16
	BA2H3	92,96	50.90	0.08
	313	92,95	50.60	0.23
	RUG2-6	92,93	51.60	0.24
	WY10	92,92	50.60	0.07
	BE23	90,14	41.10	0.51
	SH-B26	90,06	41.30	0.49
	VanAntwerpen02	83,81	27.90	2.55
B2	ATCC55673	99,28	91.70	0.03
	SW-B9	99,22	92.30	0.11
	Cu1-5	99,15	92.80	0.18
	RB-21	99,09	91.30	0.13
	NBRC 15717	97,15	85.30	0.19
	NRRL NRS-350	97,14	85.90	0.15
	ATCC 7055	97,1	85.20	0.16
	SG45	97,1	87.50	0.20
	LD79	97,1	87.70	0.19
	L2	97,08	88.90	0.12
	SG8	97,08	87.00	0.15
	TC-13	97,07	87.70	0.20
	GM	97,06	88.60	0.04
	BC-14	97,06	87.00	0.14
	BC-43	97,04	85.80	0.17
	SG53	97,03	87.00	0.15
	M5	95,49	81.70	0.23
	S4C11	95,48	78.30	0.32
	DE0175	95,35	86.00	0.20
	H1k	95,33	81.70	0.24
	OR-15	92,98	80.90	0.17
	G25-113	92,87	77.00	0.28
	juneja	85,37	57.40	0.26
	NEB1292	85,08	53.30	0.34

	ZB2	84,68	53.40	0.17
	ZC1	84,65	52.70	0.20
	1226	81,06	23.30	0.10
B3	B5	100	100.00	0.01
	DSM16288	97,97	81.60	0.32
	G25-68	87,93	34.80	1.12
B4	A3E1WT_S3	98,535	88.20	0.86
	A1E2GFP_S2	98,521	88.20	0.85
	BTU7	96,83	73.50	0.75
	p3-SID801	96,756	73.70	0.79
	MER73	96,755	73.30	0.75
	Ant232	96,755	73.60	0.94
	CK6	96,721	74.20	0.71
	2069sda1	96,684	73.60	1.07
	HMB20428	96,679	73.20	0.97
	KF3	96,649	72.80	0.75
	EC30O1N-I3-2	96,627	72.70	0.85
	EB93	96,602	73.30	0.74
	ZB201705	96,576	72.90	0.98
	CSA2	96,574	71.80	0.89
	NS1	96,493	71.20	1.09
	Q2H1	96,453	71.20	0.76
	B7	96,401	71.00	0.80
	A3E1GFP_S4	96,39	70.70	0.83
	FJAT2396	96,365	70.50	1.11
	A1E2WT_S1	96,362	70.80	0.84
	GD44	96,345	69.70	0.34
	TP802B-4	96,259	69.60	0.89
	TP10O4B-4	96,253	70.50	0.77
	44	96,224	70.10	0.86
	KF19	96,222	69.20	1.03
	SC112	96,214	69.60	0.99
	JHS1	96,169	69.40	1.06
	DSM8801	96,102	68.90	1.06
	Bi80	95,961	70.00	1.01
	3612	95,883	68.90	1.33
	242	95,436	64.40	0.83
	2RO30	93,337	52.10	0.72
	FD2	NA	22.00	0.67
	N3****	96,898	74.40	0.76
B5	B3	100	100.00	0.01
	DSM16288	97,98	81.50	0.33
	G25-68	87,95	34.80	1.11
B7	A1E2WT_S1	987.607	90.20	0.04
	A3E1GFP_S4	98713	90.20	0.03
	FJAT2396	984.922	88.20	0.31
	CSA2	977.436	81.20	0.09
	Q2H1	976.484	81.10	0.04
	SC112	975.674	79.80	0.19
	TP10O4B-4	975.521	80.20	0.03

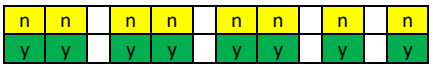
	JHS1	975.335	79.40	0.25
	TP802B-4	974.784	79.60	0.08
	3612	971.809	77.20	0.53
	Bi80	971.201	80.00	0.21
	KF3	970.253	74.80	0.05
	HMB20428	969.669	75.00	0.17
	ZB201705	969.159	75.10	0.18
	EC30O1N-I3-2	968.995	75.10	0.05
	DSM8801	968.837	74.80	0.26
	BTU7	968.421	74.10	0.05
	MER73	968.301	73.90	0.06
	GD44	967.992	72.10	0.47
	p3-SID801	967.807	74.30	0.01
	2069sda1	967.777	74.30	0.27
	Ant232	966.967	74.10	0.14
	CK6	966.754	74.40	0.09
	242	966.424	72.80	0.03
	EB93	96637	73.70	0.06
	NS1	965.554	71.90	0.29
	44	965.102	72.60	0.06
	A1E2GFP_S2	965.041	70.90	0.05
	A3E1WT_S3	964.922	70.80	0.06
	B4	96468	71.00	0.80
	KF19	96138	69.00	0.23
	2RO30	935.475	52.60	0.08
	FD2	NA	28.80	0.13
	N3****	97,024	75.30	0.04
B8	B1	100	100.00	0.00
	I4	95,6	64.10	0.14
	I6	95,27	63.30	0.20
	NBRC17520	93,5	53.00	0.44
	GGC-P6A	93,45	53.00	0.59
	DSM1321	93,4	53.10	0.35
	MGYG-HGUT-00083	93,35	53.00	0.64
	DE003	93,27	52.40	0.14
	7894	93,19	51.40	0.02
	DE0084	93,16	52.40	0.46
	CFBP13531	93,13	51.50	0.06
	MYb48	93,1	51.30	0.03
	P558	93,04	51.40	0.01
	OG2	93,01	51.10	0.16
	WY10	92,94	50.60	0.07
	RUG2-6	92,93	51.60	0.24
	313	92,92	50.70	0.23
	BA2H3	92,82	50.90	0.08
	BE23	90,15	41.10	0.51
	SH-B26	90,03	41.30	0.48
	VanAntwerpen02	83,92	27.90	2.55

Supplementary table 3, metabolisms reconstructed using KEGG mapper

B1	B8	B4	B7	B3	B5	B6	B2	
								Carbohydrate metabolism
								Central carbohydrate metabolism
y	y	y	y	y	y	y	y	M00001 Glycolysis (Embden-Meyerhof pathway), glucose => pyruvate (12)
y	y	y	y	y	y	y	y	M00002 Glycolysis, core module involving three-carbon compounds (8)
y	y	y	y	y	y	y	y	M00003 Gluconeogenesis, oxaloacetate => fructose-6P (10)
y	y	y	y	y	y	y	y	M00307 Pyruvate oxidation, pyruvate => acetyl-CoA (4)
y	y	y	y	y	y	y	y	M00009 Citrate cycle (TCA cycle, Krebs cycle) (18)
y	y	y	y	y	y	y	y	M00010 Citrate cycle, first carbon oxidation, oxaloacetate => 2-oxoglutarate (3)
y	y	y	y	y	y	y	y	M00011 Citrate cycle, second carbon oxidation, 2-oxoglutarate => oxaloacetate (15)
y	y	y	y	y	y	n	n	M00004 Pentose phosphate pathway (Pentose phosphate cycle) (8)
y	y	y	y	y	y	n	n	M00006 Pentose phosphate pathway, oxidative phase, glucose 6P => ribulose 5P (3)
y	y	y	y	y	y	y	y	M00007 Pentose phosphate pathway, non-oxidative phase, fructose 6P => ribose 5P (4)
n	n	n	n	n	n	n	n	M00580 Pentose phosphate pathway, archaea, fructose 6P => ribose 5P (2)
y	y	y	y	y	y	y	y	M00005 PRPP biosynthesis, ribose 5P => PRPP (1)
n	n	n	n	n	n	n	n	M00008 Entner-Doudoroff pathway, glucose-6P => glyceraldehyde-3P + pyruvate (3)
n	n	n	n	n	n	n	n	M00308 Semi-phosphorylative Entner-Doudoroff pathway, gluconate => glycerate-3P (4)
								Other carbohydrate metabolism
n	n	y	n	n	n	n	n	M00631 D-Galacturonate degradation (bacteria), D-galacturonate => pyruvate + D-glyceraldehyde 3P (8)
n	n	n	n	n	n	n	n	M00061 D-Glucuronate degradation, D-glucuronate => pyruvate + D-glyceraldehyde 3P (5)
n	n	y	n	y	y	y	n	M00632 Galactose degradation, Leloir pathway, galactose => alpha-D-glucose-1P (4)
y	y	y	y	y	y	y	n	M00854 Glycogen biosynthesis, glucose-1P => glycogen/starch (4)
y	y	y	y	y	y	y	n	M00549 Nucleotide sugar biosynthesis, glucose => UDP-glucose (3)
n	n	y	n	y	y	y	n	M00554 Nucleotide sugar biosynthesis, galactose => UDP-galactose (2)
y	y	y	y	y	y	y	y	M00909 UDP-N-acetyl-D-glucosamine biosynthesis, prokaryotes, glucose => UDP-GlcNAc (5)
y	y	y	y	y	y	y	y	M00012 Glyoxylate cycle (5)
n	n	n	n	n	n	n	n	M00741 Propanoyl-CoA metabolism, propanoyl-CoA => succinyl-CoA (3)
								Energy metabolism
								Carbon fixation
n	n	n	n	n	n	n	n	M00167 Reductive pentose phosphate cycle, glyceraldehyde-3P => ribulose-5P (4)
n	n	n	n	n	n	n	n	M00168 CAM (Crassulacean acid metabolism), dark (1)
y	y	y	y	y	y	y	n	M00579 Phosphate acetyltransferase-acetate kinase pathway, acetyl-CoA => acetate (2)
								Methane metabolism
n	n	y	y	n	n	n	n	M00345 Formaldehyde assimilation, ribulose monophosphate pathway (4)
								Nitrogen metabolism
y	y	y	y	y	y	n	n	M00531 Assimilatory nitrate reduction, nitrate => ammonia (2)
n	n	n	n	n	n	n	n	M00530 Dissimilatory nitrate reduction, nitrate => ammonia (2)
								Sulfur metabolism
y	y	y	y	y	y	n	y	M00176 Assimilatory sulfate reduction, sulfate => H2S (6)
								ATP synthesis
y	y	y	y	y	y	y	y	M00151 Cytochrome bc1 complex respiratory unit (3)
y	y	y	y	y	y	y	y	M00155 Cytochrome c oxidase, prokaryotes (4)
y	y	y	y	y	y	y	y	M00416 Cytochrome aa3-600 menaquinol oxidase (4)
y	y	y	y	y	y	y	y	M00157 F-type ATPase, prokaryotes and chloroplasts (8)
								Lipid metabolism
								Fatty acid metabolism
y	y	y	y	y	y	y	y	M00082 Fatty acid biosynthesis, initiation (6)
y	y	y	y	y	y	y	y	M00083 Fatty acid biosynthesis, elongation (6)
y	y	y	y	y	y	y	y	M00086 beta-Oxidation, acyl-CoA synthesis (1)
y	y	y	y	y	y	y	y	M00087 beta-Oxidation (5)
								Lipid metabolism
n	n	n	n	n	n	n	n	M00088 Ketone body biosynthesis, acetyl-CoA => acetoacetate/3-hydroxybutyrate/acetone (4)
y	y	y	y	y	y	y	y	M00093 Phosphatidylethanolamine (PE) biosynthesis, PA => PS => PE (3)
								Nucleotide metabolism
y	y	y	y	y	y	y	y	Purine metabolism
y	y	y	y	y	y	y	y	M00048 Inosine monophosphate biosynthesis, PRPP + glutamine => IMP (11)
y	y	y	y	y	y	y	y	M00049 Adenine ribonucleotide biosynthesis, IMP => ADP,ATP (5)
y	y	y	y	y	y	y	y	M00050 Guanine ribonucleotide biosynthesis, IMP => GDP,GTP (5)
								Pyrimidine metabolism
n	n	n	n	n	n	n	n	M00051 Uridine monophosphate biosynthesis, glutamine (+ PRPP) => UMP (7)
y	y	y	y	y	y	y	y	M00052 Pyrimidine ribonucleotide biosynthesis, UMP => UDP/UTP,CDP/CTP (3)
y	y	y	y	y	y	y	y	M00053 Pyrimidine deoxyribonuleotide biosynthesis, CDP => dCTP (3)
y	y	y	y	y	y	y	y	M00938 Pyrimidine deoxyribonuleotide biosynthesis, UDP => dTTP (5)
y	y	y	y	y	y	n	y	M00046 Pyrimidine degradation, uracil => beta-alanine, thymine => 3-aminoisobutanoate (4)

Amino acid metabolism

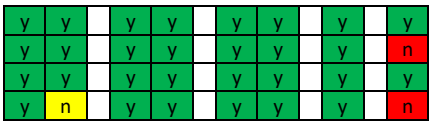
Serine and threonine metabolism



M00020 Serine biosynthesis, glycerate-3P => serine (2)

M00018 Threonine biosynthesis, aspartate => homoserine => threonine (5)

Cysteine and methionine metabolism



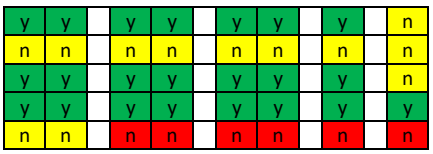
M00021 Cysteine biosynthesis, serine => cysteine (2)

M00609 Cysteine biosynthesis, methionine => cysteine (6)

M00017 Methionine biosynthesis, apartate => homoserine => methionine (9)

M00034 Methionine salvage pathway (12)

Branched-chain amino acid metabolism



M00019 Valine/isoleucine biosynthesis, pyruvate => valine / 2-oxobutanoate => isoleucine (5)

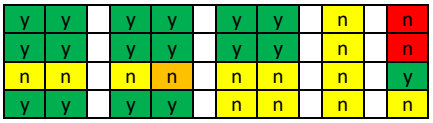
M00535 Isoleucine biosynthesis, pyruvate => 2-oxobutanoate (3)

M00570 Isoleucine biosynthesis, threonine => 2-oxobutanoate => isoleucine (6)

M00432 Leucine biosynthesis, 2-oxoisovalerate => 2-oxoisocaproate (4)

M00036 Leucine degradation, leucine => acetoacetate + acetyl-CoA (8)

Lysine metabolism



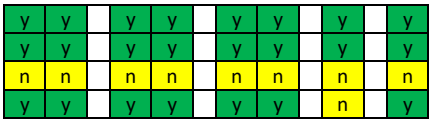
M00016 Lysine biosynthesis, succinyl-DAP pathway, aspartate => lysine (9)

M00525 Lysine biosynthesis, acetyl-DAP pathway, aspartate => lysine (9)

M00526 Lysine biosynthesis, DAP dehydrogenase pathway, aspartate => lysine (6)

M00527 Lysine biosynthesis, DAP aminotransferase pathway, aspartate => lysine (7)

Arginine and proline metabolism



M00028 Ornithine biosynthesis, glutamate => ornithine (5)

M00844 Arginine biosynthesis, ornithine => arginine (3)

M00029 Urea cycle (4)

M00015 Proline biosynthesis, glutamate => proline (3)

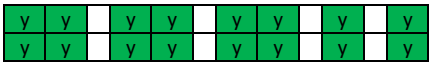
Polyamine biosynthesis

M00133 Polyamine biosynthesis, arginine => agmatine => putrescine => spermidine (4)

M00134 Polyamine biosynthesis, arginine => ornithine => putrescine (2)

M00135 GABA biosynthesis, eukaryotes, putrescine => GABA (3) (complete 3/3)

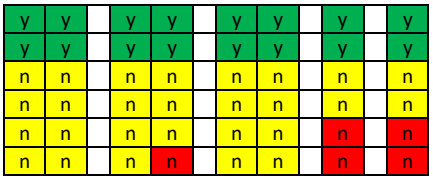
Histidine metabolism



M00026 Histidine biosynthesis, PRPP => histidine (11)

M00045 Histidine degradation, histidine => N-formiminoglutamate => glutamate (4)

Aromatic amino acid metabolism



M00022 Shikimate pathway, phosphoenolpyruvate + erythrose-4P => chorismate (9)

M00023 Tryptophan biosynthesis, chorismate => tryptophan (9)

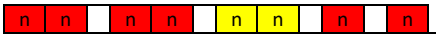
M00024 Phenylalanine biosynthesis, chorismate => phenylpyruvate => phenylalanine (3)

M00025 Tyrosine biosynthesis, chorismate => HPP => tyrosine (3)

M00533 Homoprotocatechuate degradation, homoprotocatechuate => 2-oxohept-3-enedioate (3)

M00038 Tryptophan metabolism, tryptophan => kynurenine => 2-aminomuconate (6)

Other amino acid metabolism



M00027 GABA (gamma-Aminobutyrate) shunt (3)

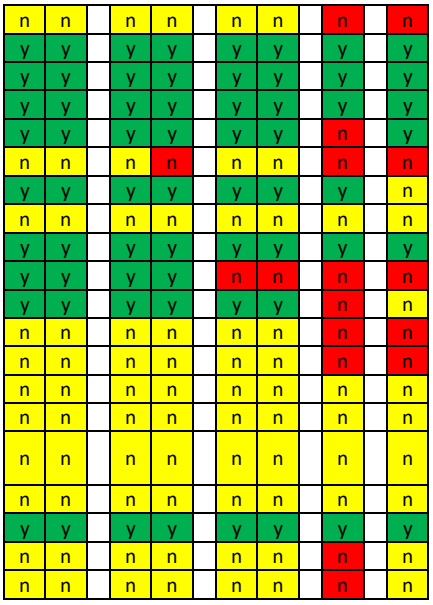
Glycan metabolism

Lipopolysaccharide metabolism

M00922 CMP-Neu5Ac biosynthesis (2)

Metabolism of cofactors and vitamins

Cofactor and vitamin metabolism



M00895 Thiamine biosynthesis, prokaryotes, AIR (+ DXP/glycine) => TMP/TPP (9)

M00899 Thiamine salvage pathway, HMP/HET => TMP (4)

M00125 Riboflavin biosynthesis, plants and bacteria, GTP => riboflavin/FMN/FAD (7)

M00916 Pyridoxal-P biosynthesis, R5P + glyceraldehyde-3P + glutamine => pyridoxal-P (2)

M00115 NAD biosynthesis, aspartate => quinolinate => NAD (6)

M00912 NAD biosynthesis, tryptophan => quinolinate => NAD (7)

M00119 Pantothenate biosynthesis, valine/L-aspartate => pantothenate (5)

M00913 Pantothenate biosynthesis, 2-oxoisovalerate/spermine => pantothenate (4)

M00120 Coenzyme A biosynthesis, pantothenate => CoA (6)

M00572 Pimeloyl-ACP biosynthesis, BioC-BioH pathway, malonyl-ACP => pimeloyl-ACP (6)

M00123 Biotin biosynthesis, pimeloyl-ACP/CoA => biotin (5)

M00573 Biotin biosynthesis, BioI pathway, long-chain-acyl-ACP => pimeloyl-ACP => biotin (4)

M00577 Biotin biosynthesis, BioW pathway, pimelate => pimeloyl-CoA => biotin (5)

M00881 Lipoic acid biosynthesis, plants and bacteria, octanoyl-ACP => dihydrolipoyl-E2/H (1)

M00882 Lipoic acid biosynthesis, eukaryotes, octanoyl-ACP => dihydrolipoyl-H (1)

M00883 Lipoic acid biosynthesis, animals and bacteria, octanoyl-ACP => dihydrolipoyl-H => dihydrolipoyl-E2 (2)

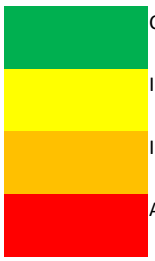
M00884 Lipoic acid biosynthesis, octanoyl-CoA => dihydrolipoyl-E2 (1)

M00126 Tetrahydrofolate biosynthesis, GTP => THF (8)

M00842 Tetrahydrobiopterin biosynthesis, GTP => BH4 (2)

M00843 L-threo-Tetrahydrobiopterin biosynthesis, GTP => L-threo-BH4 (2)

y	y	y	y	y	y	y	y	M00880 Molybdenum cofactor biosynthesis, GTP => molybdenum cofactor (6)
n	n	n	n	n	n	n	y	M00140 C1-unit interconversion, prokaryotes (3)
n	n	n	n	n	n	n	n	M00141 C1-unit interconversion, eukaryotes (1)
y	y	y	y	y	y	n	y	M00846 Siroheme biosynthesis, glutamyl-tRNA => siroheme (9)
n	n	n	n	n	n	n	n	M00121 Heme biosynthesis, plants and bacteria, glutamate => heme (10)
y	y	y	y	y	y	y	y	M00926 Heme biosynthesis, bacteria, glutamyl-tRNA => coproporphyrin III => heme (10)
y	y	y	y	n	n	n	y	M00924 Cobalamin biosynthesis, anaerobic, uroporphyrinogen III => sirohydrochlorin => cobyrinate a,c-diamide (13)
y	y	y	y	n	n	n	y	M00122 Cobalamin biosynthesis, cobyrinate a,c-diamide => cobalamin (8)
n	n	n	n	n	n	n	n	M00116 Menaquinone biosynthesis, chorismate (+ polyprenyl-PP) => menaquinol (8)
								Biosynthesis of terpenoids and polyketides
								Terpenoid backbone biosynthesis
y	y	y	y	y	y	n	n	M00096 C5 isoprenoid biosynthesis, non-mevalonate pathway (9)
y	y	y	y	y	y	n	n	M00364 C10-C20 isoprenoid biosynthesis, bacteria (2)
n	n	n	n	n	n	n	n	M00365 C10-C20 isoprenoid biosynthesis, archaea (1)
								Polyketide sugar unit biosynthesis
y	y	y	n	y	y	n	n	M00793 dTDP-L-rhamnose biosynthesis (4)
								xenobiotics biodegradation
								Aromatics degradation
y	y	y	n	y	y	n	n	M00569 Catechol meta-cleavage, catechol => acetyl-CoA / 4-methylcatechol => propanoyl-CoA (7)
								Signature modules
								Gene set
								Drug resistance
n	n	n	n	n	n	n	n	M00725 Cationic antimicrobial peptide (CAMP) resistance, dltABCD operon (5)
n	n	n	n	n	n	n	n	M00726 Cationic antimicrobial peptide (CAMP) resistance, lysyl-phosphatidylglycerol (L-PG) synthase MprF (2)
n	n	n	n	n	n	n	n	M00704 Tetracycline resistance, efflux pump Tet38 (1)
n	n	n	n	n	n	n	n	M00730 Cationic antimicrobial peptide (CAMP) resistance, VraFG transporter (3)
y	y	y	y	n	n	n	n	M00625 Methicillin resistance (3)
y	y	y	y	n	n	n	n	M00627 beta-Lactam resistance, Bla system (4)
y	y	y	y	y	y	n	y	M00700 Multidrug resistance, efflux pump AbcA (2)
n	n	n	n	n	n	n	n	M00702 Multidrug resistance, efflux pump NorB (1)
n	n	n	n	n	n	n	n	M00705 Multidrug resistance, efflux pump MepA (1)
								Module set
								Metabolic capacity
n	n	n	n	n	n	n	n	M00618 Acetogen (0)
y	y	y	y	y	y	n	n	M00615 Nitrate assimilation (1)
y	y	y	y	y	y	n	y	M00616 Sulfate-sulfur assimilation (4)



General conclusions

Thanks to the Whole Genome Sequencing techniques used in this work, it was possible to conduct more effective analyzes at the taxonomic level in comparison to the use of conventional techniques. In fact, it was possible to obtain precise classifications both for strains belonging to the *Lacticaseibacillus* spp., a group whose classification has been debated to date, and for *Bacillus* spp. strains isolated from the ancient amphora, where conventional methods, such as 16S rRNA gene sequencing due to the high similarity between species, would not have led to comparable and discriminative results.

However, several issues with the databases in use arose. The most important gene bank today is the National Center for Biotechnology Information (NCBI), which has been collecting material deposited by scientists from all over the world since 1988. Its freedom of use, while making it a critical and essential tool for global progress, does not guarantee the accuracy of the information contained therein. In fact, genomes with incorrect species identifications as well as sequencing and assembly errors have been discovered. As a result, it is necessary to verify all of the sequences used in the analyses in order to avoid the propagation of errors, which can become even more dramatic and bring negative effects on the results of research on single genomes, but also on metagenomic studies. In the future, it will be increasingly important to expand the databases available in order to make more effective comparisons. This is also true for databases used to search for genetic factors in genomes, such as virulence genes. Although effective, the tools currently available have not yet reached the level of completeness and precision required to guarantee the total safety of the aforementioned strains. Future research will be required to implement and make the available techniques even more effective and accurate.